accenture

Commonwealth of Massachusetts and Accenture LLP

AI Hub Data Commons Collaborative Consulting Services - Technical Response

RFP No. 2026-JAII-01
October 7, 2025

SUBMITTED TO:
Massachusetts Technology Collaborative
Sabrina Mansur
Director Massachusetts AI Hub

SUBMITTED BY:
Tim Rogers
Accenture LLP
781 962 9525
Timothy.P.Rogers@accenture.com

accenture

October 7, 2025

Sabrina Mansur
Director Massachusetts AI Hub
Massachusetts Technology Collaborative
Subject: Response to RFP No. 2026-JAII-01 – AI Hub Data Commons Collaborative Consulting Services

Dear Ms. Mansur,

On behalf of Accenture, we are pleased to submit our proposal to support the Massachusetts AI Hub in designing, developing, and launching the Data Commons Collaborative (DCC). The DCC is a cutting edge strategic and foundational initiative to centralize access to data for AI development, research, and innovation, and a core element of the Commonwealth's AI and economic development strategy.

The DCC is more than infrastructure—it is a commitment to shared progress, responsible innovation, and measurable public benefit. Accenture is prepared to help make that commitment real by delivering a modern, enterprise-grade data and AI platform that strengthens agency and public data sharing, informs policy through better data use, and advances equitable, secure AI-driven outcomes for Massachusetts residents.

Accenture brings two things that matter most for this work: **experience that counts and speed that matters.** We've partnered with the Commonwealth before—most recently with EOTSS to launch the AI Center of Excellence, building governance frameworks and risk processes that now guide AI adoption across 11 Executive Branch Offices. That work taught us what works here: progress through partnership, not prescription.

What the MA AI hub can expect from our work

- **Drive Smart, Scalable Progress** - We will partner with you from strategy to scale - developing strong governance, making rapid progress, and operationalizing Data & AI in ways that deliver measurable value for DCC stakeholders.
- **Bring Deep Public Sector and Local Experience** - Our team includes public service professionals with deep experience across major Massachusetts AI projects and national public and private technology programs. We understand the state's landscape, challenges, and goals, and we'll tailor every step to reflect them.
- **Best in class tools to build a secure AI Data platform** - Our experts will architect and deploy a cloud-native, high-performance platform (in partnership with AWS and Snowflake) optimized for collaboration, performance, security, and compliance to enable secure data sharing and scalable innovation.
- **Operationalize trust and governance** - Our solution is fully end-to-end and will implement policy, technical controls, monitoring, data sharing agreements, metadata standards, privacy protections, bias mitigation, and synthetic data practices to advance ethical, inclusive AI and public trust.
- **Accelerate Delivery of Tangible Value** - Our commitment is to help the MA AI Hub translate its vision into measurable value - swiftly and with confidence. Our proposed approach drives speed to value with efficiency, based on our deep experience and understanding of Massachusetts and technology.

Our proposal focuses squarely on what the DCC needs: a technically robust, socially responsible, and sustainable platform that enables rapid innovation without compromising trust. We combine proven accelerators and pragmatic steps so Massachusetts can move quickly while safeguarding privacy, equity, and security.

Thank you for the opportunity to support this important initiative. I welcome the chance to discuss our approach and next steps. Please contact me at 781 962 9525 or Timothy.P.Rogers@accenture.com

Sincerely,
Tim Rogers
Managing Director
Accenture LLP

accenture

Table of Contents

Executive Summary 4
Our Understanding 4
Proposed Solution 4
Our Approach 5
Our Proposed Team and Qualifications 7

Our Project Approach and Services 12
Project Approach 13
Phase 01 Initiation & Design 13
Phase 02 Infrastructure & Setup 14
Phase 03 Development 14
Phase 04 Testing & Governance 15
Phase 05 Rollout & Adoption 15
Services 16
Program Management & Stakeholder Coordination: 16
Technical Architecture & Platform Development 17
Dataset Curation & Integration 20
Synthetic Data Capability 22
AI Fairness & Bias Framework 23
Security, Governance & Compliance 27
Testing, Rollout & Adoption 30

Key Personnel 33

Budget & Cost Proposal 38

Relevant Experience & References 39

Disclaimer 42
Legal Disclaimer 42
Assumptions 43

Executive Summary

Our Understanding

The Massachusetts AI Hub is taking bold steps to advance its position as a leader in artificial intelligence, data innovation, and equitable technology access. This initiative is bigger than a platform, it is a shared public resource that turns the Commonwealth's data strengths into a catalyst for inclusive innovation.

We will be partners in shaping a culture of stewardship and collaboration—where participation is simple, rules are clear, and value flows back to the ecosystem. The Data Commons Collaborative (DCC) becomes a place where innovators find signal (not noise), where communities see themselves reflected in the data that shapes decisions, and where the Commonwealth's leadership in responsible AI is visible to the nation.

Guided by the AI Data Commons intent, our approach is anchored in a few enduring principles:

People at the center—design for clarity, access, and dignity in every interaction.

Trust by design—governance, transparency, security, ethical use, and accountability as defaults.

Equity as a feature—multilingual, accessible experiences and fairness as a first-order goal.

Open—and responsibly so—an extensible commons that invites participation while honoring privacy and security.

Proposed Solution

This is the work: to help Massachusetts imagine and sustain the DCC to spark discovery, accelerate ideas into impact, and keep the public good at the heart of AI. We are ready to partner with you to turn this vision into a reality—one that sets a new standard for responsible, inclusive AI innovation.

The DCC will gather and steward curated datasets, enable privacy-preserving synthetic data, and embed fairness, ethics, and security throughout so that breakthrough ideas can take root without compromising public trust.

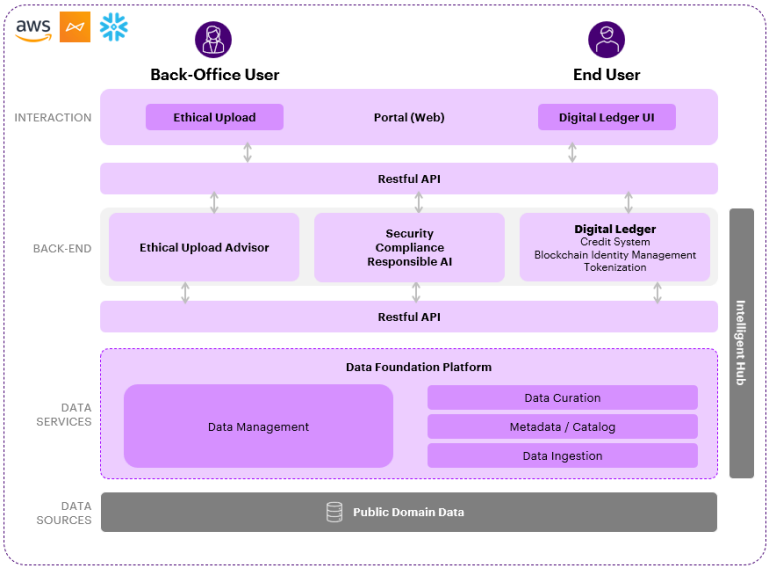
Our proposed architecture delivers the DCC as a layered platform where security, interoperability, and transparent accountability work together from the ground up.

How We'll Build It

Security by Design: **AWS** foundation with Zero Trust architecture protects every layer. Our Ethical Upload Advisor Agentic Architecture validates provenance, consent, and compliance before data enters the system—making security automatic, not optional.

Interoperability as Standard: **Snowflake and Alation** enable secure data sharing across institutions without moving sensitive information. RESTful APIs and our Intelligent Hub ensure seamless collaboration across Massachusetts' diverse ecosystem.

Conceptual Foundation Solution Architecture



Future-Ready Architecture: The Commons Credits Ledger starts with AWS QLDB for verifiable audit trails but uses adapter-based design. As needs evolve, it can migrate to Hyperledger Fabric or public blockchain anchoring—without disrupting users or requiring system redesign.

We partner with you to build a DCC where participation is simple, rules are clear, and value flows back to contributors. Where innovators find trusted data, communities see themselves reflected in decisions, and Massachusetts' leadership in responsible AI is visible nationwide.

We're ready to help Massachusetts spark discovery, accelerate ideas into impact, and keep the public good at the heart of AI.

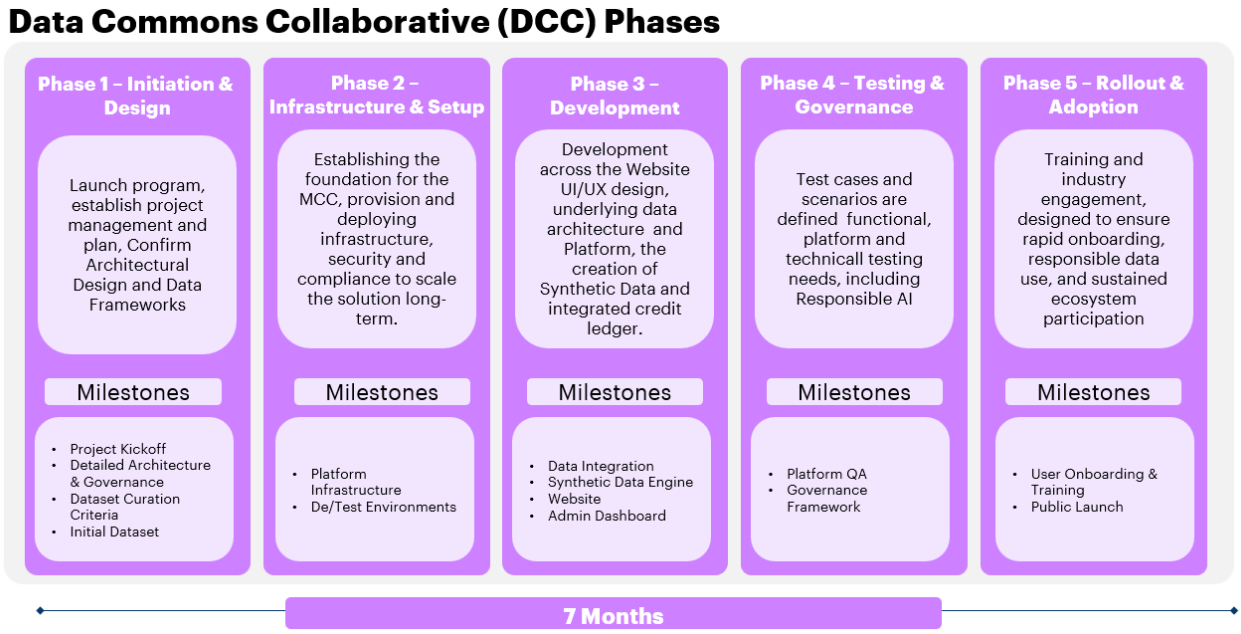
Our Approach

Grounded in the MA AI Hub's stated goals, phases and services, our proven approach to this project is designed to ensure each phase builds upon the previous one, creating a cohesive and synergistic strategy that drives the DCCs scaling and maturity.

Each phase acts as a foundational pillar, starting with establishing program governance and final design, followed by implementing the required infrastructure to enable full functionality through proven development methodologies. Testing will validate the system performance while our robust rollout framework will support a successful launch of the DCC to drive user adoption.

This approach supports the subsequent phases and ensures a seamless integration of technology, governance, and data management practices. Our proposed phases allow the MA AI

hub to **rapidly establish the platform and framework and deploy these use cases in seven months.**


Data Commons Collaborative (DCC) Phases
Phase 1 – Initiation & Design
Launch program, establish project management and plan, Confirm Architectural Design and Data Frameworks
Milestones
Project Kickoff
Detailed Architecture & Governance
Dataset Curation Criteria
Initial Dataset
Phase 2 – Infrastructure & Setup
Establishing the foundation for the MCC, provision and deploying infrastructure, security and compliance to scale the solution long-term.
Milestones
Platform Infrastructure
De/Test Environments
Phase 3 – Development
Development across the Website UI/UX design, underlying data architecture and Platform, the creation of Synthetic Data and integrated credit ledger.
Milestones
Data Integration
Synthetic Data Engine
Website
Admin Dashboard
Phase 4 – Testing & Governance
Test cases and scenarios are defined functional, platform and technicall testing needs, including Responsible AI
Milestones
Platform QA
Governance Framework
Phase 5 – Rollout & Adoption
Training and industry engagement, designed to ensure rapid onboarding, responsible data use, and sustained ecosystem participation
Milestones
User Onboarding & Training
Public Launch
7 Months

Phase 01 – Initiation & Design

We will initiate the project by working with stakeholders to build out and establish the DCC **Goals and Objectives, Requirements, Program Governance, Project Plan and Detailed Architecture and Data Governance and Catalog.**

Phase 02 – Infrastructure & Setup

In the second phase, Accenture will focus on establishing the foundation for the DCC, **provision and deploying infrastructure, security and compliance** to scale the solution long-term.

Phase 03 - Development

In the third phase, we move into the design and technical development of the DCC, orchestrating work across the **Website UI/UX design**, underlying **data architecture, Metadata Management, Ethical Data Ingestion** with **Agentic AI**, the creation of **Synthetic Data** and implementation of an **integrated credit ledger**. Our approach to development is built on a balanced approach to business process analysis, technical solutioning, and human-centered design.

Phase 04 – Testing & Governance

While we have a dedicated testing phase, testing is not a process that comes in at the end; but is part of the full development lifecycle. Test cases and scenarios are defined and linked to the corresponding features for which they test. Accenture will develop checklists, test plans and cases for **Unit Testing (UT) and System Integration Testing (SIT). User Acceptance Testing (UAT) will be planned and executed by DCC stakeholders, with Accenture providing support.** We also test for Accessibility, Penetration and Vulnerability Testing as part of our security support for the program environments. In addition, Accenture will apply the **AI Bias & Fairness Framework.**

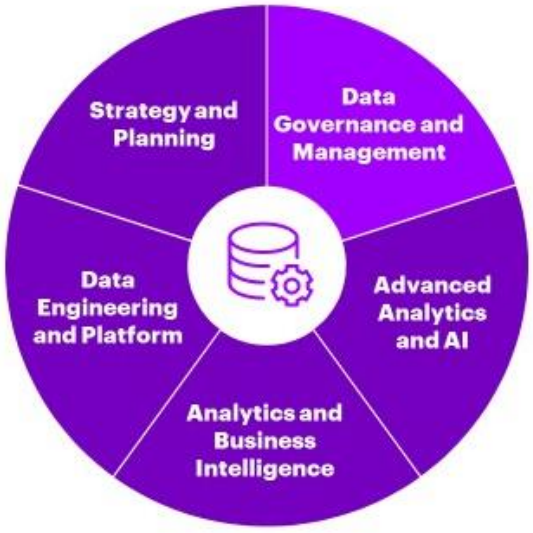
Phase 05 – Rollout & Adoption

Accenture's approach to **training and industry engagement** for the DCC is designed to ensure rapid onboarding, responsible data use, and sustained ecosystem participation. We combine a

persona-based training program with a structured engagement strategy that reflects the Commonwealth's innovation landscape and commitment to inclusive, equitable technological development.

Our Proposed Team and Qualifications

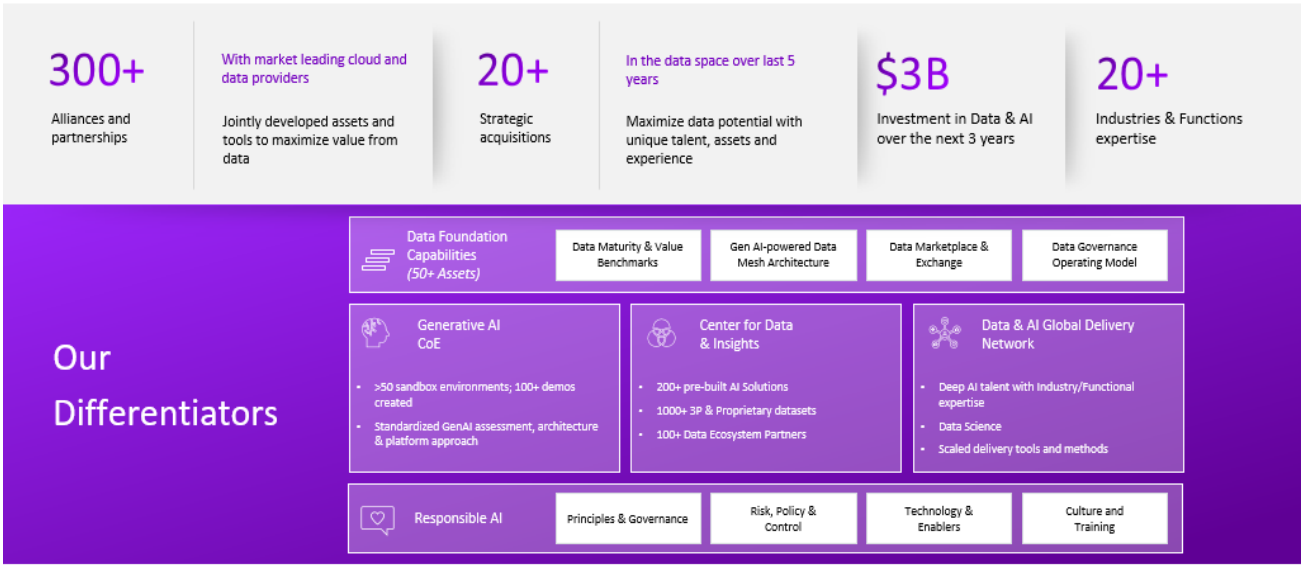
With more than 20 years of experience delivering data services for public sector clients, our Data and AI Practice is well equipped to help the MA AI Hub define, build, and scale a future-ready solution that is secure, resilient, and resident-focused. Our end-to-end capabilities – from data governance and engineering to advanced analytics and AI – are designed to evolve with the DCC's needs, supporting both immediate outcomes and long-term transformation. We understand that a modern data strategy is not just about platforms, it's about trust, interoperability, and value creation. That's why data governance and management are foundational to our approach, supported by proven frameworks aligned with the Data Management Body of Knowledge (DMBOK) principles and other leading industry standards.



Accenture offers End to End services for its clients

Our delivery model is scalable, and our specialized talent pool spans both local and national geographies – including the Commonwealth – to provide the MA AI Hub with experienced professionals who understand state government, compliance standards and local priorities.

What sets us apart is not just our technical depth—but our ability to deliver. We bring to life the data and AI strategies, we co-create with clients, transforming roadmaps into measurable


300+
Alliances and partnerships
With market leading cloud and data providers
Jointly developed assets and tools to maximize value from data
20+
Strategic acquisitions
In the data space over last 5 years
Maximize data potential with unique talent, assets and experience
$3B
Investment in Data & AI over the next 3 years
20+
Industries & Functions expertise
Our Differentiators
Data Foundation Capabilities (50+ Assets)
Data Maturity & Value Benchmarks
Gen AI-powered Data Mesh Architecture
Data Marketplace & Exchange
Data Governance Operating Model
Generative AI CoE
>50 sandbox environments; 100+ demos created
Standardized GenAI assessment, architecture & platform approach
Center for Data & Insights
200+ pre-built AI Solutions
1000+ 3P & Proprietary datasets
100+ Data Ecosystem Partners
Data & AI Global Delivery Network
Deep AI talent with Industry/Functional expertise
Data Science
Scaled delivery tools and methods
Responsible AI
Principles & Governance
Risk, Policy & Control
Technology & Enablers
Culture and Training

results. This includes integrating technology with human capital, ensuring project teams are empowered to sustain progress long after implementation.

A Leader in Data, AI Innovation and Responsible AI

Backed by a $3B investment in our data & AI practice, our mission is to help clients move from exploration to execution to real impact—responsibly and at scale. Our global network includes **72,000+ AI-skilled professionals** and **1,650+ AI-related patents**, giving Massachusetts access to the latest methods and talent. We can mobilize specialized teams quickly and right-size delivery for projects across the Commonwealth.

Snowflake — A trusted solution in the Commonwealth

With Snowflake, we run a joint adoption playbook: governance and onboarding standards, architecture and performance reviews, and clear cost transparency—so agencies can find, share, and improve governed data without friction. We're a Snowflake Elite Partner with global recognition (including Global Data Cloud Services Partner of the Year, 2024) and real depth on the ground: **9,000+ Snowflake-skilled practitioners** and **2,100+ certifications**. The effect for Massachusetts is practical—faster onboarding, safer self-service, and consistent quality across teams.

Accenture | Snowflake Partnership
Delivering Business Value through Data Cloud
#1 Snowflake Workforce
9,000+ Skilled resources
2,100+ Certified
10 trainers
Focused growth through acquisitions
Acquired and integrated strategic Snowflake consulting firms like Clarity Insights, Infinity Works, Knowledgent, Fairway, Core compete, Analyze8
Acquired Universal Data Models with 330+ industry canonical models and integrated into Snowflake verticalized offerings
Snowflake Elite Partner
Global Data Cloud Services Partner of the Year 2024
GSI Delivery Partner of Year 2023
EMEA Partner of the Year 2022
One stop shop for Data transformation providing end to end coverage – Data Integration, Business Intelligence, DataOps, Security, Governance
Verticalized solution offerings driving business value with functional use cases
Solutions & Accelerators
Clinical Study Insight Generation (CSIG) – Gen AI powered asset to help clinical researchers in collating & creating new clinical studies with chatbot, Text-to-SQL & graph capabilities.
Intelligent Data Foundation: Software asset to build repeatable, reusable and automated data engineering pipelines (data-ops) in a low code and no code way in a native Snowflake implementation leveraging Snowpark
Relevance Engine on Snowflake: Accenture's intelligence-driven solution built on Snowflake & Salesforce that enables marketers to catalyze growth & drive customer relevance by supporting key marketing activities & outcomes
MiDAS: Accenture proprietary SAP HANA to Snowflake migration automation suite, offering the ability to code convert SAP HANA objects
Operational Cockpit: Operational Governance for monitoring and optimizing queries and warehouses.
Intelligent Supply Chain on Snowflake: Suite of customizable AI-powered decision products across marketing, merchandising and supply chain functions
500+ Enterprise implementations for F500 customers
Consulting
Assessments, Cloud Data strategy, Business cases, Accelerated Transformation
Migration
Teradata, Oracle, Netezza, Hadoop, Sybase, SQL Server – object conversions, data migration
Industry solutions
Business Specific Solutions for Banking, Capital Markets, Healthcare Insurance, and Retail
Data governance
Data catalogs on SF, Data quality/Lineage, Data Security, Masking
Data Cloud
Data Sharing, Data Marketplace, Data AppDev, Analytical workloads, Snowpark
DataOps
Ingestion pipelines, CI/CD integrations, maintenance, Optimization, calibration, COEs
STRICTLY CONFIDENTIAL - ACCENTURE PROPRIETARY
Copyright © 2025 Accenture. All rights reserved. 13

Snowflake: Partnership in Practice

End-to-End Data & AI Expertise

The DCC isn't a single feature, it's an ecosystem. Users need to find trustworthy data, use it safely, improve it over time, and extend it programmatically. That means metadata, governance, APIs, analytics, and curation all working together. End-to-end expertise ensures the DCC will feel coherent and dependable from day one. Our capabilities span the full data lifecycle—from governance and engineering to advanced analytics, synthetic data, and responsible AI. We ground our approach in proven frameworks aligned with the Data Management Body of Knowledge (DMBOK) and leading industry standards, ensuring robust governance and ethical data use. A modern data strategy is about more than technology, it's about trust, interoperability, and value creation.

We are deeply embedded in the data & AI ecosystem through strategic partnerships with technology leaders, giving the MA AI Hub access to the most advanced tools, platforms, and applied innovations in the world. We also collaborate with leading academic institutions to drive cutting-edge thought leadership and ethical frameworks that support responsible AI. These partners help us shape industry standards and bring best-in-class practices to public service clients like the MA AI Hub. By tapping into this ecosystem, the Commonwealth can accelerate adoption, reduce risk, and ensure that AI solutions are transparent, equitable, and aligned with public sector values.

Strategic Partnerships Supporting Responsible AI

Real world examples of innovation in action:

- We partnered with **Amazon Web Services (AWS)** to create the Accenture Responsible AI Platform powered by AWS to help organizations adopt and scale AI rapidly with trust and confidence throughout their enterprise.
- Accenture has partnered with **Google Public Sector** to launch the Federal AI Solution Factory to accelerate the prototyping and piloting of AI-powered solutions for federal agencies.
- We work with Microsoft to provide end-to-end security solutions that reimagine cyber protection and cyber resilience solutions with generative AI including the **Accenture Adaptive Detection and Response** platform with **Microsoft Sentinel** at its core.
- Partnered to help clients leverage **NVIDIA AI Foundry and Omniverse**, paired with Accenture's deep industry expertise, for accelerated AI adoption

Core Capabilities Aligned to DCC Needs

We have identified key core capabilities we believe the DCC needs, and how Accenture is prepared to address each one.

	AI Readiness & Responsible AI Preparation of data for ethical AI, including anonymization, synthetic data generation, and embedding fairness and transparency into every solution.
	Security Robust secure development and comprehensive security controls (RBAC, MFA, encryption, audit trails) and Zero Trust principles to ensure data protection and compliance.
	Human-Centered Design & Accessibility Role-based user journeys, plain-language guidance, and WCAG-compliant accessibility to ensure the DCC is intuitive, inclusive, and easy to adopt.
	Secure Data Sharing & Governance Federated frameworks for secure, purpose-driven collaboration, plus governance templates and reusable artifacts for compliance and interoperability.
	Metadata & Data Management Robust processes to organize, classify, label, and curate data, ensuring reliable, high-quality datasets and standardized metadata management across agencies.
	Central Credits Ledger Incentive and accountability layer that tracks contributions and usage through auditable credits, encouraging sharing, improving quality, and aligning participation with program goals.

Proven Impact in similar work to the DCC initiative

We have supported similar clients to address similar challenges—fragmented data, requiring strong security requirements, and the need to move fast without compromising trust:

Client	Summary of Sample Work
USAFacts	USAFacts wanted to empower Americans with facts, not confusion. We partnered with their team to reimagine USAFacts.org, conducting deep stakeholder interviews and technical discovery to understand what citizens really need. The result? A new digital journey and design style guide that makes government data clear, actionable, and accessible—just in time for the election. The launch was met with praise from users who finally felt they could trust and use the data that shapes their communities.
City of Antwerp (Belgium)	Antwerp set out to become Europe's first true Smart City, and Accenture helped make it happen. Working with city leaders and government partners, we delivered a blockchain-based Life Event Management proof-of-concept that gives 520,000 residents control over their personal data. Our team built a solution that tracks life events securely, lets citizens share only what's needed, and makes the government a trusted certification authority. The project proved that smart cities start with smart, citizen-centered data.

Client	Summary of Sample Work
US Federal Government – CDC	The CDC needed to modernize how it managed data across 8 centers and 200+ systems. Accenture delivered the technical design and governance processes for 30+ applications, supporting a wide range of users and platforms. Our work helped the CDC streamline data management, improve compliance, and unlock new possibilities for public health innovation.
Large State Pension Agency	Preparing for the future of AI, a major state pension agency turned to Accenture to establish an AI Center of Excellence and Steering Committee. We guided them through technology readiness, business process redesign, and strategic road mapping, while developing robust architecture and security standards. Our frameworks for Responsible AI and use case prioritization helped the agency move confidently from planning to production.

Examples of Accenture's Expertise and Successful Record

What this unlocks for Massachusetts

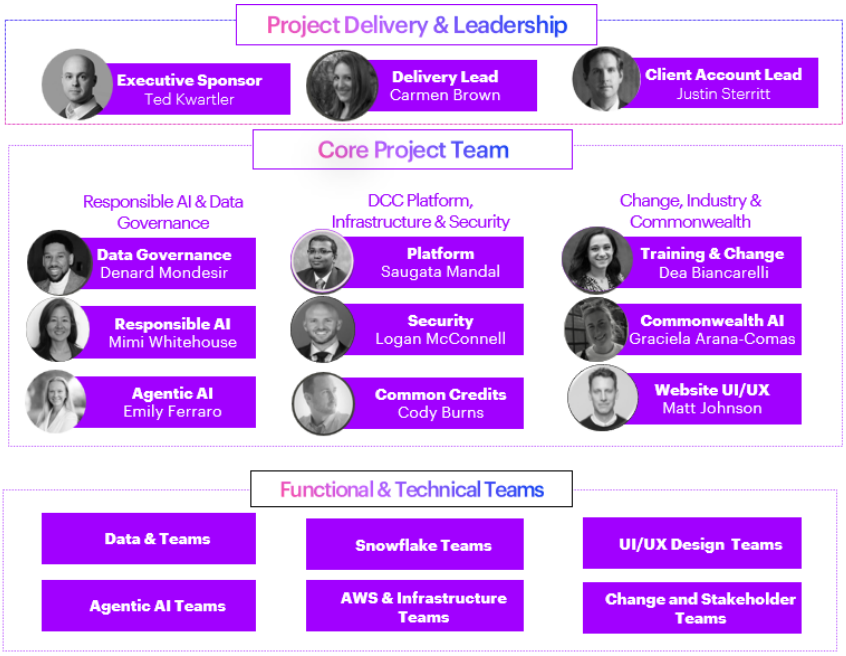
- **Speed:** Reusable marketplace patterns (access workflows, metadata automation, semantic search) shortening time-to-value.
- **Safety:** Policy/RBAC-based access with security-trimmed results, sensitive-data detection, and full auditability make responsible use, the default.
- **Sustainability:** Open, pluggable architecture (connectors, catalog integrations, semantic/graph layer) that grows with the DCC—plus training and playbooks to hand off and drive ongoing confidence in the solution.

Our Massachusetts Based Public Sector Team and Experience

We've partnered with the Commonwealth to lay the groundwork for responsible AI—most recently through the AI Center of Excellence with EOTSS. Together, we built governance frameworks, risk processes, and change strategies that help agencies adopt AI safely and confidently. That work reinforced what matters most: progress through partnership, not prescription. The DCC is the next step in that shared journey—creating a resource that serves everyone, responsibly and at scale.

We bring to the MA AI Hub DCC project, a team with a strong mix of Commonwealth talent and National expertise.

- Our Project Executive **Ted Kwartler,** and Delivery Lead, **Carmen Brown** are public and private AI experts, having driven solutions across the world.
- Our AI Lead, **Graciela Arana-Comas** and Data Lead, **Denard Mondesir** are deeply embedded with the Commonwealth and EOTSS and have each led major enterprise-wide projects in Data and AI. They know how to deliver complex programs in Massachusetts.
- Our Boston-based Change and Stakeholder team, **Dea Biancarelli** and **Justin Sterritt** led the Strategy work for the Commonwealth that supported the Governor's AI taskforce and launch of the MA AI Hub and are deeply in tune with the needs of the local business, startup and academic community.



Our Project Approach and Services

Delivering on the Commonwealth's vision requires more than technical expertise—it demands a thoughtful, future-ready approach that balances innovation, security, and inclusiveness at every step. Our methodology is designed to ensure the DCC is not only robust and scalable, but also accessible, ethical, and trusted by all stakeholders.

To achieve this, we have provided our **detailed project approach** to the essential DCC **services**:

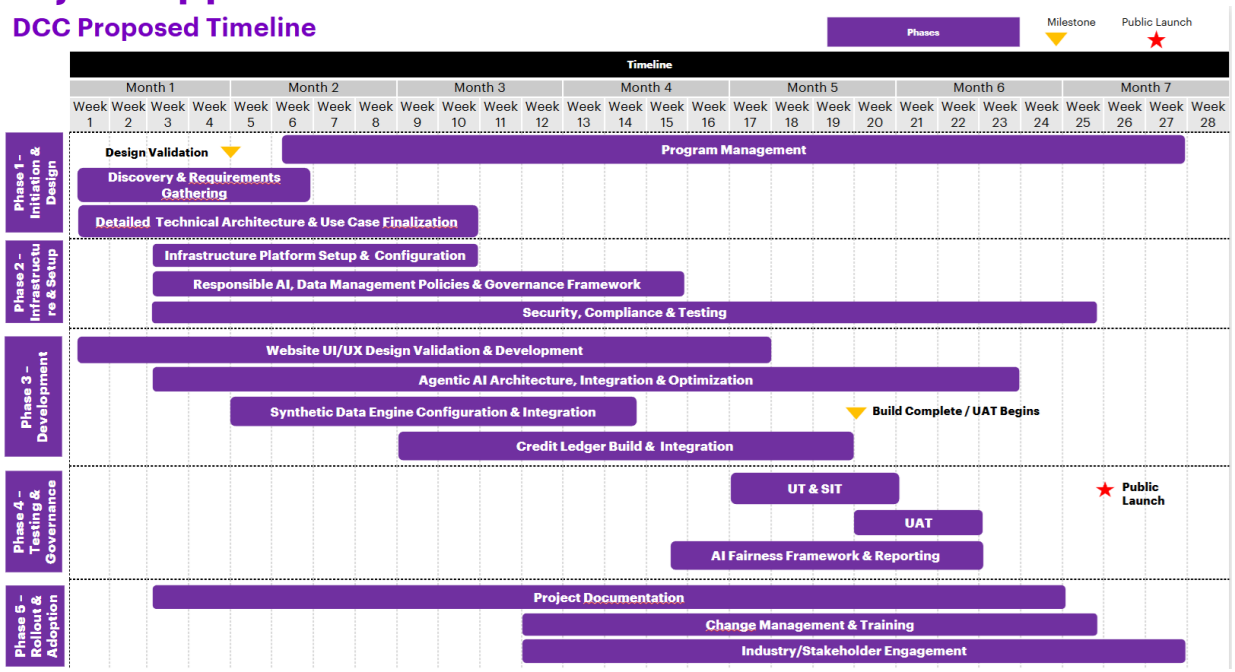
- **Program Management & Stakeholder Coordination**
- **Technical Architecture & Platform Development**

- **Data Curation & Ingestion**
- **Synthetic Data Capability**
- **AI Fairness & Bias Framework**
- **Security, Governance & Compliance**
- **Testing, Rollout & Adoption**

Each of these **services** is essential to building the DCC so that it will be able adapt to evolving needs, empower diverse users, and set a new standard for responsible AI in Massachusetts. The following sections detail how we will bring these services to life through five key **phases – Initiation and Design, Infrastructure & Setup, Development, Testing & Governance, and Rollout & Adoption.**

Project Approach

DCC Proposed Timeline


Phases
Milestone
Public Launch
Timeline
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Week 1 Week 2 Week 3 Week 4 Week 5 Week 6 Week 7 Week 8 Week 9 Week 10 Week 11 Week 12 Week 13 Week 14 Week 15 Week 16 Week 17 Week 18 Week 19 Week 20 Week 21 Week 22 Week 23 Week 24 Week 25 Week 26 Week 27 Week 28
Phase 1 – Initiation & Design
Phase 2 – Infrastructure & Setup
Phase 3 – Development
Phase 4 – Testing & Governance
Phase 5 – Rollout & Adoption
Design Validation
Program Management
Discovery & Requirements Gathering
Detailed Technical Architecture & Use Case Finalization
Infrastructure Platform Setup & Configuration
Responsible AI, Data Management Policies & Governance Framework
Security, Compliance & Testing
Website UI/UX Design Validation & Development
Agentic AI Architecture, Integration & Optimization
Synthetic Data Engine Configuration & Integration
Build Complete / UAT Begins
Credit Ledger Build & Integration
UT & SIT
Public Launch
UAT
AI Fairness Framework & Reporting
Project Documentation
Change Management & Training
Industry/Stakeholder Engagement

Phase 01

Initiation & Design

We will initiate the project by working with stakeholders to build out and establish the DCC Goals and Objectives, Requirements, Program Governance, Project Plan and Detailed Architecture and Data Governance and Catalog.

Our team will leverage our extensive experience in developing and implementing data, analytics, and AI policy requirements and governance frameworks for the MA AI Hub, both in the United States and internationally. We will develop a detailed roadmap outlining phases, milestones, and timelines for the formation of the DCC. This includes creating a governance committee structure to oversee the DCC, guiding data, analytics, and AI with a cohesive strategy and vision. By clearly

defining roles, responsibilities, and operating procedures, we will ensure the sustainability and effectiveness of the DCC.

We will also use this phase as an opportunity to engage key DCC and Commonwealth stakeholders in state government, the startup and business community, and the academic community to truly define the goals, objectives, vision and use cases. We have built relationships with many of these stakeholders through our work with the EOTSS and the greater business and academic community based on our deep technical experience working with them.

Once our operating and governance structures are agreed, we will begin discovery activities to ensure our technical architecture aligns to business goals. Co-creative workshops with stakeholders and end users will help to set the stage for high-level experience and technical design concepts. With the support of business analysts, technologists, and user-experience designers, we will align on program vision and definitions for success; we will describe various user-experience needs across diverse user personas and mindsets; and we will document end-to-end user journeys.

Building on these initiation and design phase activities will allow Accenture and the MA AI Hub to finalize the technical architecture which will provide the foundational effort and cost necessary for the subsequent phases and full system buildout. We will also use this phase to establish the DCC initial data governance and data catalog standards to set the path for data ingestion and use.

Phase 02

Infrastructure & Setup

In the second phase, Accenture will focus on establishing the foundation for the DCC, provisioning and deploying infrastructure, security and compliance, to scale the solution long-term.

Our team will work with our partners, Snowflake, AWS, and Alation to implement the baseline instances and deploy the infrastructure necessary to establish the DCC. We will also refine our data governance and responsible AI framework so the infrastructure is embedded with these principles from the start.

Phase 03

Development

In the third phase, we will move into the technical development of the DCC, orchestrating work across the Website UI/UX design, underlying data architecture, Metadata Management, Streamlined Ethical Data ingestion with Agentic AI, the creation of Synthetic Data and implementation of an integrated credit ledger.

Our partnership with Snowflake and AWS for this solution will leverage Snowflake's unified analytics platform to enable secure, scalable, and high-performance data. Our team will combine deep expertise in cloud-native architecture and data engineering with AWS's advanced cloud

capabilities and Snowflake's powerful data technology to deliver a robust platform. This system will support scheduled data ingestion, advanced analytics, and AI-driven insights, all while ensuring compliance with stringent regulatory and privacy requirements.

Together with Snowflake and AWS, we will deliver a modern data solution that empowers the DCC to transform its data into actionable intelligence. This system will promote interoperability, discoverability, and privacy-preserving data sharing, advancing the Commonwealth Data and AI maturity and enabling impactful, data-driven solutions to business problems, academic research and public services.

In addition to implementation of technology platforms, in phase 3, we will design user-experience patterns, design systems, and user interfaces. To maximize the adoption and efficacy of technical solutions, Accenture uses human-centered design methods to understand exactly how user-needs and behaviors influence solution requirements. We create intuitive, useful, and personalized digital tools and experiences by continuing to engage with both business stakeholders and end-users.

As in Phase 1, co-creative activities like user interviews and stakeholder workshops continue to help us avoid biased or assumptive decision-making in the design of both UX and UI. We also use cycles for design experimentation and iteration to translate co-creative insight into flexible and scalable design systems that support diverse and divergent user-needs. Our approach to design also embraces WCAG 2.2 AA accessibility standards, which are foundational to not only our design explorations, but are also a part of our usability testing and design refinement process.

Phase 04

Testing & Governance

While we have a dedicated testing phase, testing is not a process that comes in at the end; but is part of the full development lifecycle. Test cases and scenarios are defined and linked to the corresponding features for which they test. Accenture develops checklists, test plans and cases for Unit Testing (UT), System Integration Testing (SIT), and User Acceptance Testing (UAT). We also will complete Accessibility, Penetration and Vulnerability Testing as part of our overall support for the program environments.

In addition to testing the system for the necessary technical and functional considerations, there will be dedicated testing through the application of the AI Bias & Fairness Framework to ensure the responsible AI and data governance framework have been implemented. This will help ensure that at launch, the approach to reviewing the underlying data for the AI use cases will be tested for both functional need and identification of bias so that mitigation techniques can be applied.

Phase 05

Rollout & Adoption

Accenture's approach to training and industry engagement for the DCC is designed to ensure rapid onboarding, responsible data use, and sustained ecosystem participation. We combine a persona-based training program with a structured engagement strategy that reflects the

Commonwealth's innovation landscape and commitment to inclusive, equitable technological development.

Our plan delivers operational readiness for internal teams, guided enablement for external users, and strategic outreach to industry partners—aligning with the DCC's milestones for User Onboarding & Training Delivered and Public Launch. All materials will be published on a searchable knowledge base and designed to meet WCAG accessibility standards. Completion metrics, time-to-first dataset/API call, and feedback loops will be tracked to measure adoption and improve content iteratively.

Services

Program Management & Stakeholder Coordination:

Accenture's approach to program management is grounded in decades of experience delivering complex, multi-phase initiatives for public sector clients. For the DCC, we will bring a disciplined, transparent, and agile coordination model that spans all phases of delivery—from onboarding and training to public launch and ecosystem engagement.

Our PMO serves as the central coordination hub, ensuring that all workstreams operate under a unified plan of record. This includes maintaining consistent delivery rhythms, enforcing quality and governance controls, and enabling real-time visibility for stakeholders across agencies, academia, and industry.

Working Cadence and Controls

We operate with a predictable and collaborative cadence that keeps teams aligned and stakeholders engaged:

- **Weekly stand-ups** provide tactical coordination across workstreams, surfacing blockers and aligning priorities.
- **Periodic demos** showcase real user journeys through the "front-door" portal—allowing stakeholders to see how users discover, request, use, and contribute data.
- **Monthly Steering Committee meetings** offer strategic oversight, reviewing progress against milestones, managing risks, and guiding decisions.

To maintain delivery discipline, we implement a suite of governance tools:

- **Architecture Decision Records (ADRs)** document key technical choices and ensure traceability.
- **Open risk and issue logs** are actively managed and reviewed to address dependencies and mitigate delivery risks.
- **Quality gates** are enforced at key stages to validate readiness and completeness.
- **Change control processes** ensure that any scope adjustments are reviewed and approved without compromising schedule or budget.

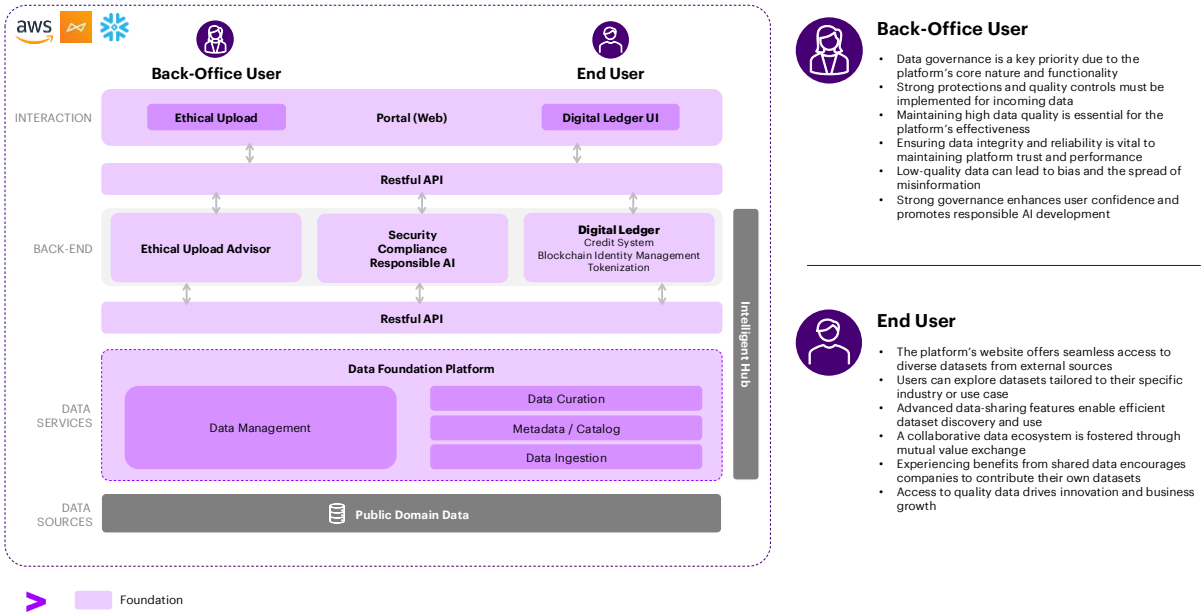
This structure ensures that all phases of the DCC program move forward in lockstep, with clear accountability and shared visibility.

Stakeholder Engagement and Communication
Our PMO ensures that stakeholders remain informed and engaged throughout the program lifecycle. The **front-door portal** is used as a demonstration environment, allowing stakeholders to experience real user flows early and often. This builds trust and provides meaningful opportunities for feedback.

We maintain a **single integrated workplan** across all workstreams, updated weekly and reviewed during governance meetings. This ensures that all teams, from technical delivery to training and outreach, are aligned on timelines, dependencies, and outcomes.

Technical Architecture & Platform Development

Conceptual Foundation Solution Architecture


aws
Back-Office User
End User
INTERACTION
Ethical Upload
Portal (Web)
Digital Ledger UI
Restful API
BACK-END
Ethical Upload Advisor
Security
Compliance
Responsible AI
Digital Ledger
Credit System
Blockchain Identity Management
Tokenization
Restful API
Intelligent Hub
Data Foundation Platform
DATA SERVICES
Data Management
Data Curation
Metadata / Catalog
Data Ingestion
DATA SOURCES
Public Domain Data
Foundation
Back-Office User
Data governance is a key priority due to the platform's core nature and functionality
Strong protections and quality controls must be implemented for incoming data
Maintaining high data quality is essential for the platform's effectiveness
Ensuring data integrity and reliability is vital to maintaining platform trust and performance
Low-quality data can lead to bias and the spread of misinformation
Strong governance enhances user confidence and promotes responsible AI development
End User
The platform's website offers seamless access to diverse datasets from external sources
Users can explore datasets tailored to their specific industry or use case
Advanced data-sharing features enable efficient dataset discovery and use
A collaborative data ecosystem is fostered through mutual value exchange
Experiencing benefits from shared data encourages companies to contribute their own datasets
Access to quality data drives innovation and business growth

The DCC will be architected to deliver a secure, scalable, and ethically grounded environment for data collaboration and innovation. The following section details the technical architecture and platform development priorities, structured around five key areas: **website and backend infrastructure, metadata management and discovery, role-based access and identity, central ledger tracking, and API-enabled data access with usage analytics**, all summarized in the integrated platform workflow section at the bottom of this section.

Data Commons Website and Backend Infrastructure
The DCC website will serve as the primary user interface, designed for intuitive navigation and rich interactivity. Built using modern web frameworks and containerized for scalability, the front-

end connects seamlessly to a robust backend infrastructure hosted on AWS. Core services are deployed across isolated VPCs, leveraging AWS Control Tower for multi-account provisioning and management. Security is foundational: all assets are encrypted at rest and in transit with AWS KMS, and network segmentation ensures sensitive operations remain on private subnets. Continuous auditing is performed through AWS CloudTrail and Macie, while GuardDuty and Security Hub monitor for anomalous or unauthorized activity across the environment.

Metadata Management, Cataloging, and Discovery

Effective data governance and discoverability are achieved through an integrated metadata management stack. Our proposal focuses on Alation as the core enterprise data catalog platform. Alation serves as the searchable index of all data assets across S3 and Snowflake, embedding governance and compliance policies into the workflow. Metadata tags highlight sensitive information, approved use cases, or retention requirements, creating a governance-by-design model. This architecture enables both raw and curated datasets to be cataloged with rich, policy-driven metadata, including provenance, consent, sensitivity classification, and terms of use. Our proposed Ethical Upload Advisor Agentic Architecture acts as a programmable enforcement point in the data ingestion pipeline, verifying technical validity and ethical integrity before metadata is generated and pushed downstream. This ensures every dataset is discoverable with full transparency into its origins, constraints, and compliance status.

Role-Based Access and Identity Management

We will employ a comprehensive role-based access control (RBAC) system integrated with federated identity providers to support secure, fine-grained permissions management. User identities are authenticated via SAML or OAuth2, enabling seamless collaboration across multiple institutions while maintaining strict access boundaries. Data access policies are enforced at both the infrastructure and application layers, leveraging Snowflake's built-in capabilities for secure sharing and fine-grained controls. This ensures that users only see and interact with datasets they are authorized to access, supporting confidential research and regulatory compliance.

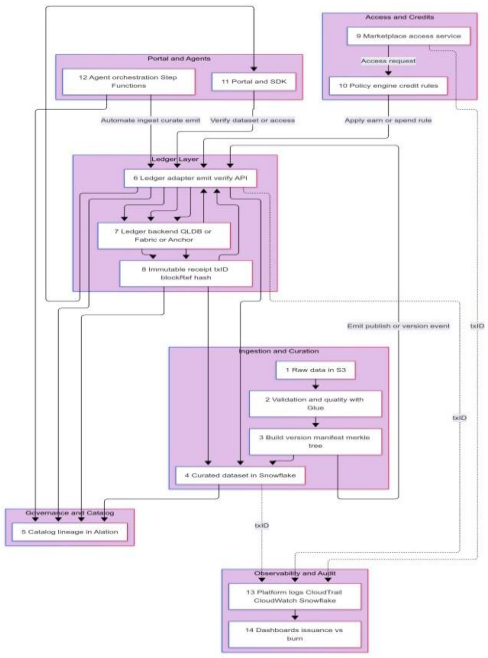
Central Ledger for Tracking

A central ledger system is implemented to provide a tamper-evident, auditable record of all significant actions within the platform. This samples ledge (figure below) captures data uploads, access requests, consent verifications, and any changes to dataset metadata or governance policies. By integrating with both the Ethical Upload Advisor Agentic Architecture and the metadata catalog, the ledger ensures a persistent chain of trust and accountability—facilitating compliance with frameworks such as HIPAA, GDPR, and evolving regional standards. All ledger entries are time-stamped and cryptographically secured, enabling transparent review by both internal auditors and external stakeholders.

Approach & Key Considerations: Credit Ledger

CONCEPTUAL PROVENANCE TRACKING FLOW

	#	What happens	Layer in the narrative view
INGEST & STORE	1	Raw data lands in S3	Ingestion and curation on AWS landing zone
	2	Glue validates schema and quality	Ingestion pipelines with automated checks
	3	Version manifest and merkle tree built	PROV-inspired event prep with hashing
	4	Curated data stored in Snowflake	Snowflake as the collaboration plane
	5	Catalog and lineage in Alation	Governance and discovery
	6	Ledger adapter emits or verifies events	Stable emit verify API, pluggable layer
	7	Ledger backend writes immutable entry	QLDB or Fabric or anchor under adapter
	8	Receipt with txID and hash returned	Minimal on chain, proofs off chain
REQUEST & TRACK	9	Marketplace requests access	Access and grant flows
	10	Policy engine applies earn spend rules	Credits economics encoded as policy objects
	11	Portal verifies dataset or access	One click verify UX and SDKs
	12	Agent orchestrates end to end steps	Agentic workflows with mediated writes
	13	txID stitched into logs and queries	CloudTrail CloudWatch Snowflake logs
	14	Dashboards show issuance burn costs	Program health and transparency


Access and Credits
9 Marketplace access service
Access request
10 Policy engine credit rules
Portal and Agents
12 Agent orchestration Step Functions
11 Portal and SDK
Automate ingest curate emit
Verify dataset or access
Apply earn or spend rule
Ledger Layer
6 Ledger adapter emit verify API
7 Ledger backend QLDB or Fabric or Anchor
8 Immutable receipt txID blockRef hash
Emit publish or version event
txID
Ingestion and Curation
1 Raw data in S3
2 Validation and quality with Glue
3 Build version manifest merkle tree
4 Curated dataset in Snowflake
txID
Governance and Catalog
5 Catalog lineage in Alation
txID
Observability and Audit
13 Platform logs CloudTrail CloudWatch Snowflake
14 Dashboards issuance vs burn

API-Enabled Data Access with Usage Analytics

To support programmatic workflows and interoperability, the DCC will expose a suite of secure, documented APIs for data access, metadata queries, and platform integration. These APIs are governed by the same RBAC policies and are monitored for usage patterns and anomalous access attempts. Usage analytics are collected and visualized in real time, providing insights into data demand, user engagement, and operational efficiency. This data will be used to optimize platform performance, inform governance decisions, and demonstrate value to stakeholders.

Integrated Platform Workflow

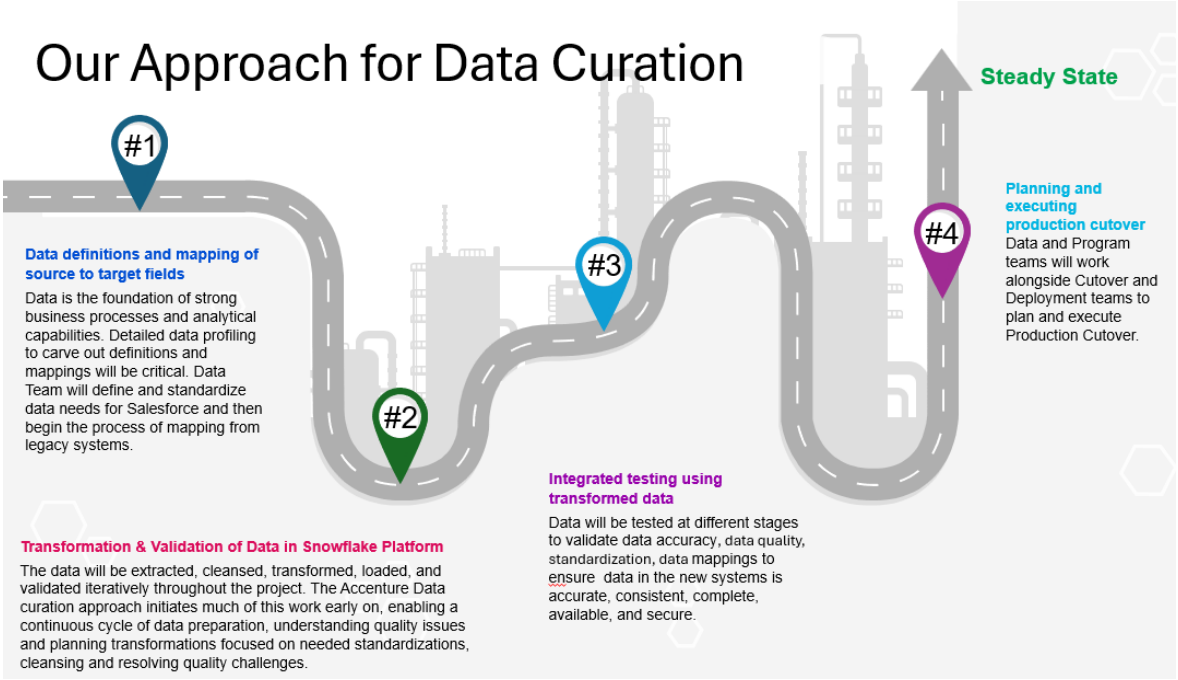
From data ingestion to discovery and usage, the platform will operate as a cohesive ecosystem. The Ethical Upload Advisor Agentic Architecture enforces both technical and ethical validation as data lands in secure S3 buckets, integrating with AWS-native services for event logging and compliance checks. Only validated data proceeds through normalization pipelines and is stored in Snowflake, where it becomes accessible for research and collaboration. Metadata and provenance are continuously carried forward, cataloged by Alation, and surfaced through the website and APIs. All actions are tracked in the central ledger, ensuring transparency and traceability throughout the data lifecycle.

This architecture ensures that the DCC not only meets the highest standards of security, compliance, and usability, but also provides a flexible foundation for future growth and evolving policy requirements. By integrating robust infrastructure, advanced metadata management, secure access controls, transparent tracking, and analytics, the DCC platform will empower stakeholders to collaborate confidently and responsibly.

Dataset Curation & Integration

The DCC platform will only be as strong, as innovative and as effective the Data that powers it. To that end we have proposed a robust methodology for Data curation and Integration withing the system. The image below shows our 4-step path to Data Curation-

1. **Data Definitions and Mapping**
2. **Transformation & Validation of Data in Snowflake**
3. **Integrated Testing**
4. **Planning and Executing Production control**


Our Approach for Data Curation
Steady State
#1
Data definitions and mapping of source to target fields
Data is the foundation of strong business processes and analytical capabilities. Detailed data profiling to carve out definitions and mappings will be critical. Data Team will define and standardize data needs for Salesforce and then begin the process of mapping from legacy systems.
#2
Transformation & Validation of Data in Snowflake Platform
The data will be extracted, cleansed, transformed, loaded, and validated iteratively throughout the project. The Accenture Data curation approach initiates much of this work early on, enabling a continuous cycle of data preparation, understanding quality issues and planning transformations focused on needed standardizations, cleansing and resolving quality challenges.
#3
Integrated testing using transformed data
Data will be tested at different stages to validate data accuracy, data quality, standardization, data mappings to ensure data in the new systems is accurate, consistent, complete, available, and secure.
#4
Planning and executing production cutover
Data and Program teams will work alongside Cutover and Deployment teams to plan and execute Production Cutover.

Ensuring High-Quality, Ethical, and FAIR Data for the DCC

The DCC will establish a robust foundation for collaborative research and innovation by prioritizing systematic curation and seamless ingestion of datasets. Our approach will support the requirement that every dataset, whether natural or synthetic, is discoverable, trustworthy, and compliant with technical and defined ethical standards.

Assisting in Curating the Initial Set of Datasets

Curation will begin with the careful selection and preparation of the initial datasets. Leveraging both manual expertise and AI-driven workflows, we will transform raw, siloed data into organized assets. The Ethical Upload Advisor Agentic Architecture acts as the first line of defense, validating that each dataset meets both technical validity and ethical integrity thresholds before inclusion. This process includes schema validation, privacy checks, and compliance with institutional and regulatory requirements. Every dataset is cataloged in Alation, ensuring that all data—natural or synthetic—receives equal governance and oversight.

Ensuring Metadata Completeness and Adherence to FAIR Principles

A cornerstone of our curation strategy is the comprehensive enrichment of metadata. Each dataset is accompanied by detailed, standardized metadata to ensure it is Findable, Accessible, Interoperable, and Reusable (FAIR). Metadata completeness is continuously monitored and

enforced throughout the data lifecycle, enabling researchers and stakeholders to easily discover and understand available datasets. Cataloging in Alation and integration with a unified React-based portal provides a single point of access, enhancing transparency and facilitating informed decision-making.

Supporting Ingestion Pipelines, Version Control, and Provenance Tagging

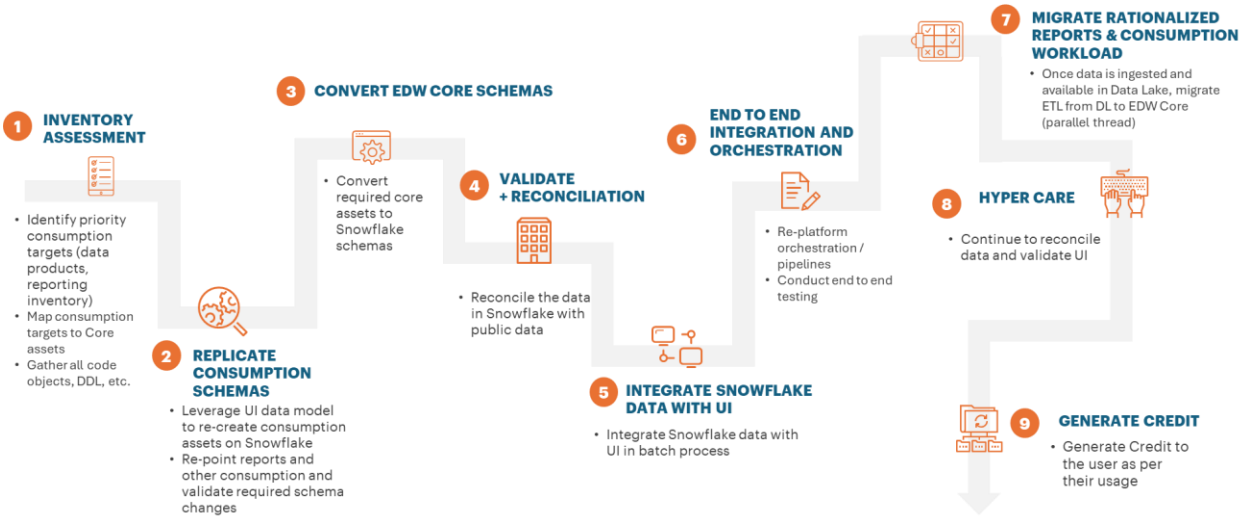
The ingestion of datasets is streamlined through automated pipelines built on AWS SageMaker, EKS, and supporting DevOps tools like Terraform and AWS CDK. These pipelines are equipped with built-in versioning and provenance tagging, ensuring every dataset's origin, transformation history, and update chronology are immutably logged. The central credits ledger initially implemented via AWS QLDB, records every dataset upload and transformation into a cryptographically verifiable journal, supporting transparency and auditable workflows. This ledger can evolve to leverage permissioned blockchain technologies, such as Hyperledger Fabric, for future multi-party governance scenarios.

Automation and Agentic Orchestration

Agentic workflows extend automation across the first part of the curation and ingestion process. For example, AI agents monitor S3 for new data, trigger the Ethical Upload Advisor Agentic Architecture, start the normalization and validation of datasets, start the process of metadata enrichment, and provide needed information to generate synthetic datasets as needed, and log key pertaining actions automatically. This approach reduces manual intervention, increases scalability, and ensures that every process remains repeatable, secure, and compliant.

By integrating comprehensive curation practices, rigorous metadata management, and advanced ingestion pipelines, the DCC will transform data management from a static repository into a dynamic, governed ecosystem. Every dataset is not only ready for immediate use but is also positioned to support long-term research, innovation, and public service delivery.

Supporting the end-to-end DCC workflow is a challenging endeavor, requiring security, integrations, credits and much more. We have proposed a high-level schematic (figure below) to demonstrate that end-to-end process.


1 INVENTORY ASSESSMENT
• Identify priority consumption targets (data products, reporting inventory)
• Map consumption targets to Core assets
• Gather all code objects, DDL, etc.
2 REPLICATE CONSUMPTION SCHEMAS
• Leverage UI data model to re-create consumption assets on Snowflake
• Re-point reports and other consumption and validate required schema changes
3 CONVERT EDW CORE SCHEMAS
• Convert required core assets to Snowflake schemas
4 VALIDATE + RECONCILIATION
• Reconcile the data in Snowflake with public data
5 INTEGRATE SNOWFLAKE DATA WITH UI
• Integrate Snowflake data with UI in batch process
6 END TO END INTEGRATION AND ORCHESTRATION
• Re-platform orchestration / pipelines
• Conduct end to end testing
7 MIGRATE RATIONALIZED REPORTS & CONSUMPTION WORKLOAD
• Once data is ingested and available in Data Lake, migrate ETL from DL to EDW Core (parallel thread)
8 HYPER CARE
• Continue to reconcile data and validate UI
9 GENERATE CREDIT
• Generate Credit to the user as per their usage

Synthetic Data Capability

Enabling Scalable, Secure, and Compliant Data Innovation

The DCC will integrate synthetic data generation as a foundational capability to address privacy, compliance, and data accessibility challenges. By embedding robust synthetic data pipelines and leveraging state-of-the-art tools, the DCC will ensure that high-quality, privacy-preserving datasets are readily available for research, analytics, and innovation—without compromising the integrity or confidentiality of sensitive information.

Integration of Synthetic Data Tools

To deliver scalable and flexible synthetic data solutions, the DCC will be capable of incorporating leading platforms such as Synthea or CARLA. These tools will be seamlessly integrated into automated data ingestion and curation pipelines built on AWS SageMaker and EKS, with orchestration managed through DevOps frameworks like Terraform and AWS CDK. This automated approach ensures that synthetic data can be generated on-demand, tailored to the requirements of various analytical and machine learning use cases.

- Synthea: Specializes in creating realistic healthcare records, supporting public health research while safeguarding patient privacy.
- CARLA: Facilitates the creation of autonomous driving simulation data, crucial for AI and safety model development.

Automated Synthetic Dataset Generation & Validation Pipelines

Agentic workflows extend automation across the synthetic data lifecycle. AI agents monitor data sources for new or updated datasets, trigger synthetic data generation when privacy or access constraints are detected, and validate outputs using built-in quality and compliance checks. Enrichment and normalization are performed automatically, ensuring that synthetic datasets are representative, statistically robust, and aligned with real-world distributions.

All synthetic data activities are transparently logged in the central credits ledger, providing cryptographically verifiable records of dataset creation, transformation, and validation. This comprehensive audit trail supports reproducibility, regulatory compliance, and cross-agency collaboration.

Compliance with Privacy and Audit Policies

The DCC's synthetic data pipelines will be engineered with privacy and compliance at their core. The Ethical Upload Advisor Agentic Architecture will be invoked to assess risk, enforce privacy-preserving transformations, and validate that synthetic datasets cannot be reverse-engineered to reveal sensitive information. Automated provenance tagging and versioning ensure that every synthetic dataset's lineage is immutably recorded, supporting internal and external audits.

- Privacy-by-design: Synthetic datasets are generated using methods that prevent re-identification and ensure regulatory compliance (e.g., HIPAA, GDPR).
- Auditability: Every action—generation, validation, access—is recorded in a cryptographically verifiable journal, enabling end-to-end traceability.
- Governance: Synthetic data workflows are governed by permissioned access controls, metadata tagging, and continuous monitoring for policy adherence.

Benefits and Impact

By integrating advanced synthetic data capabilities, the DCC will transform sensitive or restricted datasets into safe, high-utility resources for research and innovation. Consumers can accelerate data sharing and analytics initiatives, reduce compliance risk, and support a culture of responsible data stewardship.

- Accelerated access to high-quality, privacy-compliant data for analytics and AI development
- Reduced manual intervention and improved scalability through agentic orchestration and automation
- Transparent, auditable processes supporting regulatory and ethical requirements
- Enhanced cross-agency and public-private collaboration via safe data sharing mechanisms

Synthetic data capability is a cornerstone of the DCC's commitment to secure, ethical, and innovative data management. Through the integration of best-in-class tools, automated pipelines, and rigorous compliance frameworks, the DCC empowers stakeholders to unlock the full value of data while upholding the highest standards of privacy and trust.

AI Fairness & Bias Framework

The DCC recognizes the critical importance of fairness, transparency, and accountability in the development and deployment of AI systems. As agencies and stakeholders increasingly rely on data-driven decision-making, the risk of perpetuating or amplifying existing biases in datasets and algorithms becomes a central concern. To address these challenges, the DCC will include a comprehensive AI Fairness & Bias Framework, designed to support synthetic and real-world datasets, as well as models derived from them, to meet defined standards of integrity and inclusiveness.

The approach laid out for the development of an AI Fairness and Bias Framework is based on the following **key considerations**:

- **The need for fairness & bias tools includes not only models, but also GenAI solutions**: The open-source tools listed in the table in section 12.7 of the RFP, such as Fairlearn and AI Fairness 360, enable fairness evaluation for predictive AI models. However, solutions for GenAI solutions are also needed.
- **Customization for fairness analyses may be necessary to be effective**: Accenture has performed fairness analyses that are customized to different types of models, and these analyses require demographic data.
- **Determination of bias depends on the definition of fairness**: Fairness can be defined in terms of parity in model results across demographic groups, or equal or equitable representation of subgroups within datasets, among other definitions. However, the applicability of different fairness definitions depends on the data available for comparison by subgroup.

Developing an AI Risk Management & Control Framework *example*

As part of a larger ask to develop a foundational program to operationalize the RAI principles, Accenture implemented a fairness framework and toolkit to standardize fairness assessments and bias mitigation practices for data scientists developing AI models.

CLIENT RESULTS:

- Defined and operationalized a foundational Responsible AI program roadmap tailored to the client's needs and priorities.
- Implemented an RAI governance structure utilizing the risk categorization approach and risk management and control framework.
- Implemented fairness framework and toolkit to standardize fairness assessment and bias mitigation practices for data scientists developing AI models.
- Developed a comprehensive RAI training program including the definition of personas and tailored learning plans.
- Client demonstrated commitment to their Responsible AI principles and put them into practice across the organization,.

- **Bias depends on the use case**: Ultimately, the intended use of datasets and models depends on the use case, and in some applications, bias is desirable. For example, it may be desirable for the purpose of reaching specific demographics like targeted communications to lower income constituents regarding services.
- **Operational bias is beyond the scope of this project**: Bias can enter into AI applications not only through data and models, but also in ongoing operations. All three components work together and can reinforce each other. For example, data returned to the DCC may contain bias. However, assessing bias within the DCC operations is outside the scope of this proposed phase.

Proposed Approach

Our approach to providing a fairness & bias framework seeks to address bias proactively, addressing and reviewing for bias in data and models on the portal.

Initially, we will review regulatory and voluntary AI governance frameworks (e.g., NIST AI Risk Management Framework and Generative AI Profile), and the DCC principles (human-centric, trustworthy, inclusive, innovative, sustainable) against which the fairness & bias framework and metrics will be developed. We will work with the AI Hub team members to determine which aspects of these frameworks are most important to include in fairness evaluations and will select metrics based on those priorities.

Once we have narrowed down the AI Hub's priorities in its assessment goals, we will develop a customized fairness assessment framework for data and models that we describe the types of

bias all evaluations will investigate. In the sections below, we have outlined proposals for data and model bias assessment workflows, though we will adjust these workflows based on the unique needs of the AI Hub. The steps in the bias assessment workflows will inform which metrics we will make available within the portal.

Next, we will select a list of metrics, like demographic parity, that could be used across use cases and with different AI models; development of bias assessments; and development of customized metrics that leverage open-source libraries. We will explore the various tools and methods provided by third party organizations, including from sources like AI Fairness 360, Fairlearn, AWS SageMaker Clarify, Hugging Face evaluate, and DeepEval. We will work with the AI Hub team to select a list of metrics for inclusion in the portal. The metrics we select will also depend on the balance of generative versus predictive models available on the portal.

We propose applying these metrics to five (5) high priority datasets to start so that the project team can have the time to review outputs, document the results in a dataset card, and review metrics in a dashboard (e.g., Streamlit dashboard) for review together to receive feedback and tune any metrics. If datasets are not available on the portal by the time the team is ready to do the fairness assessment, we will apply the evaluation to publicly available datasets. We will coordinate with the AI Hub team on selecting models to evaluate, as our priority will be models that are available on the portal. We have assumed that model creation will be at a later stage; should model assessment be required during this project period, we would be happy to discuss inclusion.

Proposed Data Bias Assessment Framework

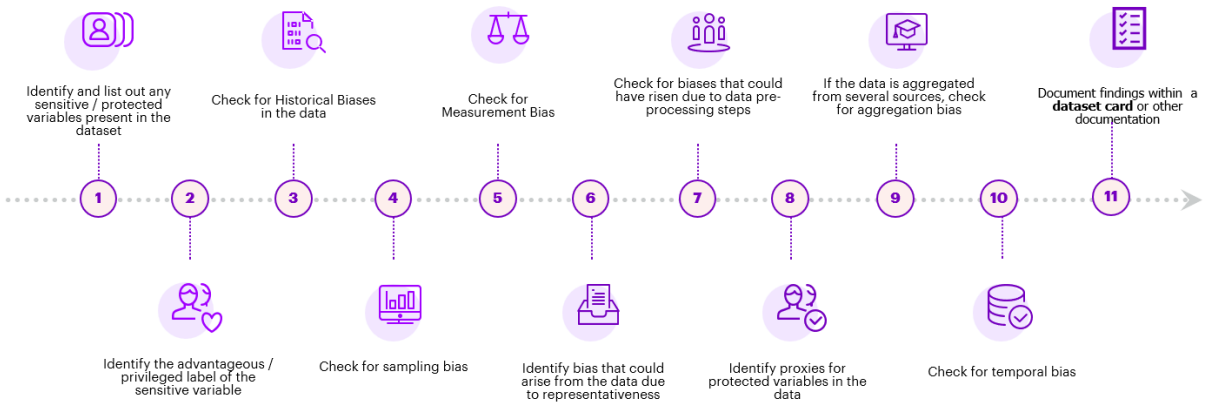
	Type	Description
DATA COLLECTION	**Historical Bias**	Historical bias is the already existing bias and socio-technical issues in the world and can seep into from the data generation process, even given a perfect sampling and feature selection
	Sampling Bias	Occurs when the training data is not collected randomly but specific dynamics or characteristics influence the fact that a data point will be collected in the data
	Measurement Bias	Bias appearing from the way of measuring/choosing/utilizing features and how we draw conclusions from them
DATA PREPARATION	**Representation Bias**	When not all subgroups are represented equally in input data
	Proxy Bias	Using proxies as a replacement for a variable to get a particular result, while that proxy is not representative for the variable
	Pre-processing Bias	Arises during data pre-processing in model development, when an operation (e.g., missing value treatment, data cleansing, outlier treatment, encoding, scaling or data transformations for unstructured data, etc.) causes or contributes to systematic disadvantage
	Aggregation Bias	Aggregation bias is an assumption that the mapping from inputs to labels is consistent across groups, whereas in reality, this is not the case
	Temporal Bias	Arises from differences in populations, systems, and behaviors over time

Data bias can come from data collection as well as during data preparation

In the table above, we include types of bias commonly found in datasets. For example, a type of data bias – sampling bias – can occur when data is collected from responses to online questionnaires or is scraped from social media. The datasets on which results are based are comprised of samples that are neither randomized nor representative of a population other than

the users of a particular online platform. Such datasets are not generalizable yet frequently used to train ML applications which are deployed for use in broader socio-technical contexts, even though data representing certain societal groups may be excluded. As part of the fairness evaluation, we will identify the subset of bias types to evaluate within the fairness evaluation based on AI Hub priorities, and then we will conduct that evaluation on the selected datasets.

Below are our initial proposed steps for conducting a dataset bias assessment. These steps would be reviewed and updated as the team works more closely with project stakeholders. Note that once bias is identified, we would record our findings within a dataset card, metadata tag, or another document type to accompany the dataset within the portal.


Identify and list out any sensitive / protected variables present in the dataset
Check for Historical Biases in the data
Check for Measurement Bias
Check for biases that could have risen due to data pre-processing steps
If the data is aggregated from several sources, check for aggregation bias
Document findings within a dataset card or other documentation
1
2
3
4
5
6
7
8
9
10
11
Identify the advantageous / privileged label of the sensitive variable
Check for sampling bias
Identify bias that could arise from the data due to representativeness
Identify proxies for protected variables in the data
Check for temporal bias

Proposed steps for data bias assessment

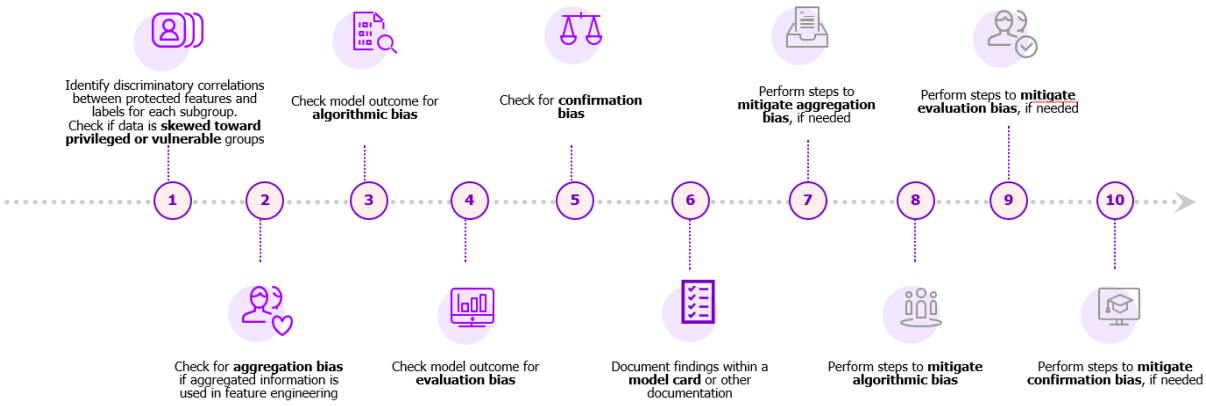
Proposed Model Bias Assessment Framework

Multiple types of bias can be identified within machine and deep learning models.

Type	Description
Aggregation Bias	Aggregation bias during model development happens when false assumptions or generalizations that are drawn for a subgroup based on observing other different subgroups or entire population is used in model training
Algorithmic Bias	Bias introduced by the choice of algorithm architecture, parameters, or optimization functions, which causes the model to create unfair or subjective outcomes.
Evaluation Bias	Arises when the testing populations do not equally represent the various groups of the user population, hence the model performance on different granularity is not consistent.
Confirmation Bias	The tendency to process information by looking for, or interpreting, information that is consistent with one's existing beliefs.

Understanding model bias types

Much like with data biases noted above, we propose steps for reviewing models. These steps would be reviewed and updated as the team works more closely with project stakeholders. We will also review and update these steps depending on the portfolio of models being assessed, as some of these evaluations are not applicable to generative AI models due to opacity and access constraints of these models. Finally, we will document our findings using a method that aligns with the portal's structure (e.g., model cards) and will update assessment reports as model metric results change. Note that the team will develop the framework for evaluation. Applying this framework to models is proposed for a future phase.


Identify discriminatory correlations between protected features and labels for each subgroup. Check if data is **skewed toward privileged or vulnerable** groups
Check model outcome for **algorithmic bias**
Check for **confirmation bias**
Perform steps to **mitigate aggregation bias**, if needed
Perform steps to **mitigate evaluation bias**, if needed
1
2
3
4
5
6
7
8
9
10
Check for **aggregation bias** if aggregated information is used in feature engineering
Check model outcome for **evaluation bias**
Document findings within a **model card** or other documentation
Perform steps to **mitigate algorithmic bias**
Perform steps to **mitigate confirmation bias**, if needed

Potential steps for review of model bias during the development phase

Taken together, the dataset and model bias review will support the DCC's goals of equity, transparency, and adherence to ethical standards in alignment with stated principles of human-centric, trustworthy, inclusive, and innovative data and AI development.

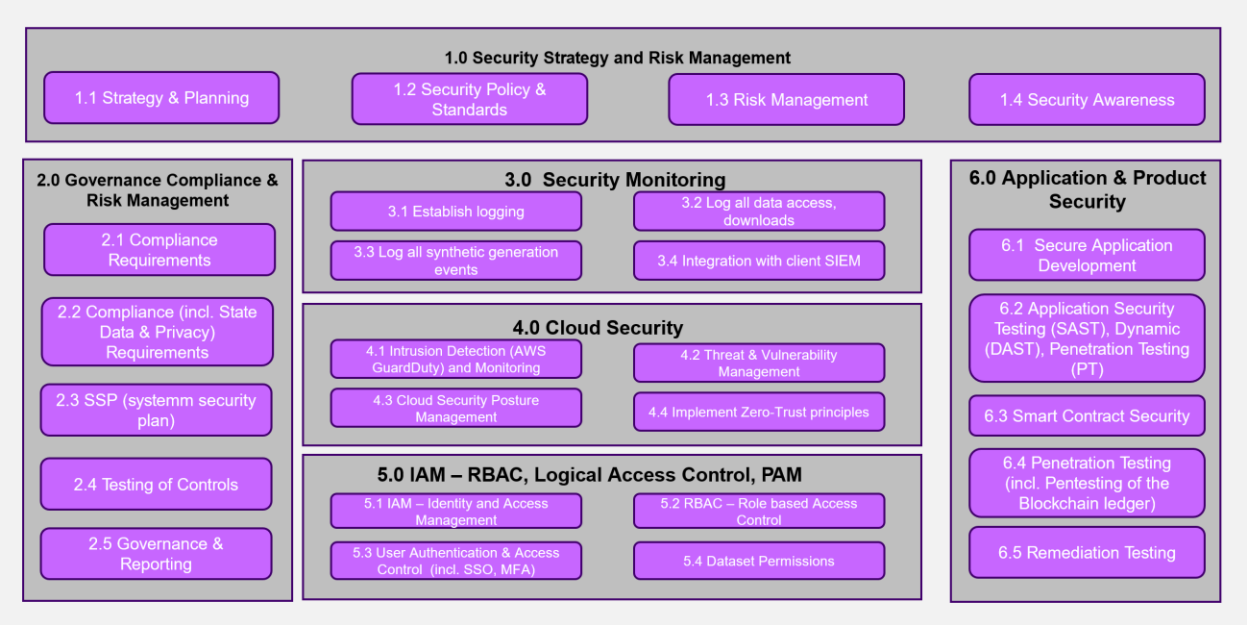
Security, Governance & Compliance

Our proposed approach is grounded in a methodology that goes beyond meeting immediate requirements; we aim to establish a sustainable, future-proof security posture. Our solution is fundamentally secure by design, ensuring security is a core element of your architecture, not a bolted-on component. This entire methodology is built on a comprehensive framework explicitly engineered to address the DCC's needs across compliance, cloud setup, data security, IAM, and application security. By strictly aligning our efforts to Zero Trust principles, we will establish continuous verification and least-privilege access across your environment, creating a unified and resilient security fabric designed for the modern threat landscape.

Zero Trust Approach

Our proposed approach is grounded in a methodology that goes beyond meeting immediate requirements; we aim to establish a sustainable, future-proof security posture. Our solution is fundamentally secure by design, ensuring security is a core element of your architecture, not an afterthought.

Security Framework


1.0 Security Strategy and Risk Management
1.1 Strategy & Planning
1.2 Security Policy & Standards
1.3 Risk Management
1.4 Security Awareness
2.0 Governance Compliance & Risk Management
2.1 Compliance Requirements
2.2 Compliance (incl. State Data & Privacy) Requirements
2.3 SSP (systemm security plan)
2.4 Testing of Controls
2.5 Governance & Reporting
3.0 Security Monitoring
3.1 Establish logging
3.2 Log all data access, downloads
3.3 Log all synthetic generation events
3.4 Integration with client SIEM
4.0 Cloud Security
4.1 Intrusion Detection (AWS GuardDuty) and Monitoring
4.2 Threat & Vulnerability Management
4.3 Cloud Security Posture Management
4.4 Implement Zero-Trust principles
5.0 IAM – RBAC, Logical Access Control, PAM
5.1 IAM – Identity and Access Management
5.2 RBAC – Role based Access Control
5.3 User Authentication & Access Control (incl. SSO, MFA)
5.4 Dataset Permissions
6.0 Application & Product Security
6.1 Secure Application Development
6.2 Application Security Testing (SAST), Dynamic (DAST), Penetration Testing (PT)
6.3 Smart Contract Security
6.4 Penetration Testing (incl. Pentesting of the Blockchain ledger)
6.5 Remediation Testing

Accenture security framework provides privacy and security by design to achieve information assurance.

Our security by design approach has the following workstreams:

Governance and Compliance

Accenture will work closely with the key stakeholders and vendors to establish security best practices to govern the DCC. Key components of governance and compliance best practices that Accenture will establish include:

- Defining processes and policies (driving security policies by consumerization of cloud services)
- Establishing the foundation for an effective audit and compliance program by providing (i) a robust System Security Plan (SSP) documenting the controls, (ii) the procedures to assess, maintain, and update that plan; along with (iii) a template for Plans of Action and Milestones (POAMs)
- Adhering to the applicable NIST standards (NIST Cybersecurity Framework, Risk Management Framework, and SP 800-53, SP 800-171, amongst others)
- Establishing a Cybersecurity Management Plan (CMP)

Cloud Security

We propose to build a secure cloud environment based on zero trust architecture principles, by leveraging the Accenture Secure Cloud Foundation (SCF) assets. The key features of SCF include:

- Cloud Security Quick Start - Establish a template-based, risk-aligned foundational architecture including a secure landing zone and roadmap for baseline cloud security.
- Security Automation - Guardrails are automatically deployed.
- Proactive Compliance - Self-healing functions continuously enforce policies in alignment with regulations and standards.
- Operations Optimization - Leverage cloud native tools and integrate with the existing SIEM and CSPM tools, to manage the cloud cost-effectively and at scale.

The Accenture Secure Cloud Factory asset includes a library of secure cloud governance policies, and "secure-from-start" provisioning templates.

Identity Access Management (IAM)

Accenture's proposed solution will implement identity and access management through AWS IAM, which will federate with the state IAM as well as with Snowflake.

Accenture will use the RBAC approach to control user access, assigning permissions based only on users' roles rather than directly. Accenture will create and implement an RBAC model that considers business and Agency roles, job functions, and individual responsibilities. Oversight of role assignment and maintenance will be ensured through a robust RBAC governance process. We will follow zero trust principles to ensure all services and resources are protected from unauthorized use. Accenture is thought leader is zero trust architecture and we have several mature assets we will bring to implement Zero Trust Architecture (ZTA) for the DCC environment

Application and Product Security

Accenture's application security strategy follows DevSecOps principles to confirm that security best practices are integrated within DevOps processes. Establishing a 'Security as Code' culture will provide flexibility and collaboration between release engineers and the security team. Tools and frameworks such as Veracode SAST, SCA, and dynamic code scanning provide identification and remediation of security risks as early as possible in the release cycles.

We will design and build a secure blockchain. This will include (i) secure development (static/dynamic analysis), (ii) secure key management, (iii) secure access through RBAC, MFA, and least privilege, (iv) Zero-Trust node communication via TLS 1.3 mutual authentication, firewall segmentation, and Intrusion Detection, and (v) Robust API security.

API Security will be rigorously tested. At a high level, the following best practices and controls are followed to ensure API security:

- Static and dynamic application security testing (SAST/DAST) for APIs
- Enforcing least privilege principle by limiting access to API functionalities
- Rotation of keys
- Measures to prevent unauthorized access
- Third-party applications have limited data access and actions, as approved by the customer
- Rate limiting

Penetration testing will be performed for each release to verify that the solution, the associated APIs and the blockchain ledger are fully secured.

Data Encryption

Accenture utilizes industry leading data encryption methods and standards, including encrypting data at rest as well as in-transit. Our solution utilizes data encryption to encrypt all data at rest using AES-256 data encryption standard. All communication is encrypted using TLS 1.3, which provides transport layer security.

Logging, Encryption, and Audit Trails

Comprehensive logging is integrated across all platform functions, capturing user activities, data access events, and system changes in real time. These logs are securely stored and encrypted, leveraging industry-standard protocols to protect both data at rest and in transit. Audit trails are automatically generated, supporting transparent and immutable records for all critical actions within the data and AI pipelines. This auditability not only facilitates internal reviews but also ensures readiness for external regulatory inspections, reinforcing the DCC's commitment to accountability and traceability. Regular reviews of logs and audit trails can be conducted to detect anomalies, support incident response, and drive continuous improvement of security practices.

Terms of Use and Enforce Ethical Data Handling

To uphold ethical standards and regulatory compliance, the DCC will have clear terms of use established, governing all data and platform interactions. These terms explicitly address permissible use, data privacy, and security obligations, as well as protocols for reporting and remediating violations. Enforcement mechanisms are embedded within workflows, leveraging automated monitoring and policy checks to ensure consistent adherence to ethical guidelines. Training and guidance are provided to both end users and administrators, fostering a culture of responsible data stewardship. The DCC's governance model will incorporate metadata tagging and policy monitoring, ensuring that ethical considerations are embedded throughout the data lifecycle. This approach supports scalable, repeatable compliance and reinforces trust among stakeholders, partners, and the broader public.

Testing, Rollout & Adoption

We have proposed a robust testing and rollout strategy to ensure a seamless transition from implementation to operational excellence.

Testing and Rollout

Testing is not a process that comes in at the end; but is part of the full development lifecycle. Test cases and scenarios will be defined and linked to the corresponding features for which they test. Accenture develops checklists, test plans and cases for the following tests:

- Unit Testing: Unit testing is developed and performed by the developer against each feature developed. A feature will not move on to the next phase of testing without successfully passing unit tests and having configuration or custom code reviewed by Technical Leads
- System Integration Testing (SIT): This type of testing brings together multiple components and features in a system configuration to execute system and integration testing demonstrating an Epic or consolidation of User Stories. Depending on the features

implemented in the sprint, the system integration testing focuses on how both internal and external components integrate to support the required features.
- User Acceptance Testing (UAT): The objective of User Acceptance Testing is to have business users and end users test the system to ensure that it is in line with requirement expectations. This typically involves a subset of testing scenarios to expose business and end users to major areas of the system for system acceptance as well as driving more ad-hoc real-world user testing that sometimes test cases don't anticipate. We recommend conducting UAT using converted masked data to provide as realistic testing as possible.

Accenture will work with relevant third parties and DCC stakeholders to perform the following test phases:

- Accessibility Testing: Accessibility compliance is a critical part of the State's solutions. Accessibility testing executes test scenarios to support application usage for persons with disabilities in line with State accessibility standards and guidelines. If appropriate, Accenture will work with State disability agencies and/or 3rd party Accessibility vendors for validation.
- Penetration and Vulnerability Testing: Accenture will work, as needed, with State technical resources for the execution of vulnerability testing and address identified issues and concerns.

Environments: To support the required development operations and the activities described above, Accenture is proposing to stand up the following environments:

- Development/Testing Environment: Used for all sprint development, unit testing, and system integration testing.
- QA Environment: Used for User Acceptance Testing, Accessibility Testing and Vulnerability Testing, as well as an environment to support training and other parallel development activities when needed, such as Data Migration development and testing. Scaled similar to production environment capacity, including a close to full set of de-identified production data.
- Production Environment: Environment solely used for production application usage.

Training and Industry Engagement

Accenture's approach to training and industry engagement for the DCC will be designed to ensure rapid onboarding, responsible data use, and sustained ecosystem participation. We will combine a persona-based training program with a structured engagement strategy that reflects the Commonwealth's innovation landscape and commitment to inclusive, equitable technological development.

Our plan delivers operational readiness for internal teams, guided enablement for external users, and strategic outreach to industry partners, aligning with the DCC's milestones for **User Onboarding & Training Delivered** and **Public Launch**.

Training Program

We will deliver a comprehensive, role-based training program aligned to the DCC governance roles and user journeys. Training content will be tailored to five key personas:

- **Admins & Operators**: Platform operations, governance controls, and incident response.

- **Curators & Stewards**: Metadata standards, publication workflows, and bias/fairness review.
- **Researchers & Developers**: Dataset discovery, access requests, API usage, and synthetic data tagging.
- **Partners & Contributors**: Submission workflows, licensing, metadata requirements, and credit visibility.
- **Auditors & Reviewers**: Access patterns, audit trails, lineage, and reporting.

Training modalities will include:

- Quick start guides and cheat sheets
- Micro-lessons (≤7 min) embedded in the portal
- Guided tours and tooltips
- Hands-on labs (Jupyter/SQL)
- Office hours and a cross-sector change-champion network

All materials will be published on a searchable knowledge base and designed to meet WCAG accessibility standards. Completion metrics, time-to-first dataset/API call, and feedback loops will be tracked to measure adoption and improve content iteratively.

Industry Engagement Strategy

Accenture brings a proven track record of driving adoption across complex stakeholder ecosystems. Our engagement strategy for the DCC will be grounded in understanding stakeholder value, building trust, and enabling co-development.

Stakeholder Discovery

We will lead structured interviews across priority sectors like healthcare, climate, transportation, and manufacturing to identify data needs, barriers to participation, and opportunities for innovation. These insights will inform tailored use cases that demonstrate how the DCC participation can accelerate research, reduce operational costs, and expand market access.

Change Champion Network

To support cultural alignment and adoption, we will establish a Change Champion Network composed of representatives from key industry verticals. These champions will:

- Serve as the liaison between their organizations and the DCC program team
- Provide feedback on platform features, governance policies, and onboarding processes
- Drive internal awareness and alignment within their organizations

This network will be modeled on Accenture's enterprise change management frameworks, successfully deployed across large-scale public and private sector transformations.

Governance and Co-Development

We will implement a structured engagement model that includes:

- Regular stakeholder roundtables, working groups, and advisory sessions
- Transparent governance processes for data access, consent, and ethical use
- Co-development opportunities for standards, tools, and pilot projects

This supports governance that is not only compliant and secure, but also participatory and responsive to community input.

Communications and Outreach
To build momentum and visibility, we will launch a multi-channel communications campaign leveraging:

- Industry consortia and trade associations
- Conferences, webinars, and innovation summits
- Media partnerships and thought leadership content
- Targeted outreach to underrepresented and emerging players in the AI ecosystem

We will use trusted engagement vehicles, including academic institutions, civic tech groups, and startup accelerators to ensure broad representation and sustained participation.

Key Personnel

Summary

Accenture has assembled a project team of highly experienced professionals who excel in data engineering, analytics, cloud platforms, systems integration, Centers of Excellence, data culture, governance, change management, and operations. We recognize that this initiative is not simply a technology implementation—it represents a fundamental shift in how Massachusetts state agencies collaborate, share insights, and leverage data as a strategic asset.

The success of the DCC will depend not only on technical excellence, but also on thoughtful change management, proactive stakeholder engagement, and a deep understanding of the Commonwealth's vision and needs. That's why we've brought together subject matter advisors and leadership with a blend of advanced technical expertise and deep public sector experience—including team members who have delivered similar solutions for Massachusetts and peer states.

Our team's dedication to customer-centric service means we anticipate and respond to your needs throughout the project lifecycle. We maintain high standards through a rigorous quality assurance process, and we are committed to knowledge transfer and capacity building to ensure lasting value for the Commonwealth. We know that when the right people come together with a shared purpose, real change becomes possible.

Organization Chart

We have organized our proposed project team to bring the technical and functional expertise in an efficient and effective manner. Our team is also primarily based in Massachusetts and has served the Commonwealth, local businesses and startups and local universities extensively over the last 25 years.

Our team is led by **Ted Kwartler, Carmen Brown and Justin Sterritt** who serve different key roles. Ted is a national leader in deployment of cutting-edge public and private AI and data solutions. Carmen is a local project executive, deeply experienced with orchestrating the technical delivery of large-scale data projects in the Commonwealth. Justin is a local account lead and former public official focused on ensuring MA AI Hub and the DCC success and engagement with Massachusetts stakeholders in the public, private and academic sectors.

The core project team on the program will each be responsible for a variety of tasks and workstreams, overseen daily by Carmen. Each brings a unique and essential skill, from Massachusetts Data governance expertise like **Denard Modenisr**, **Graciela Arana-Comas** in AI excellence or robust Snowflake platform experience like **Saugata Mandal**. The Core project team will interact with their respective teams, across AI, Data, Platform and Design to effectively deliver the project.

Proposed Key Staff

Effective leadership is crucial for team success. We bring experienced technology leaders with relevant industry knowledge to collaborate with your teams, understand your requirements, and deliver solutions tailored to your objectives. In the table below, we outline the qualifications of our personnel who will serve in key leadership roles and work with you and your teams.

Ted Kwartler
Project Executive

Ted Kwartler is an experienced strategist in AI and data transformation, recognized for turning ambition into executable roadmaps. He has led enterprise data programs and Responsible AI frameworks across complex organizations, aligning governance, funding, and delivery. Ted is known for sequencing use cases so value shows up early and scales, and for partnering with executives to translate strategy into results teams can run—consistently and at pace. **Ted is a Core Local Member of the Accenture Massachusetts Team.**

Justin Sterritt
Client Account Lead

Justin is a seasoned public service consultant and former Massachusetts government executive, bringing decades of experience in strategic planning, financial management, and stakeholder engagement. As the former Chief of Administration and Finance and CFO for the City of Boston, he led citywide initiatives and managed senior stakeholders. Justin has also held senior finance roles in the Massachusetts State Legislature and Executive Branch, navigating complex policy and regulatory challenges to drive outcomes. At Accenture, he oversees public service work across state and local government and higher education, ensuring solutions are tailored to client needs. **Justin is a Core Local Member of the Accenture Massachusetts Team.**

Carmen Brown
Delivery Lead

Carmen is a Senior Manager in Accenture Technology, who has led major technology transformation programs across education, healthcare, energy, finance, and public sector clients, with deep expertise in complex delivery, functional architecture, and performance test engineering. She has successfully directed large-scale platform upgrades, digital customer experience programs, and compliance-driven policy initiatives, managing diverse teams of developers, testers, designers, and analysts. Carmen is known for aligning strategy and execution, driving end-to-end delivery across the full lifecycle, and consistently achieving outcomes that improve scalability, customer engagement, and regulatory compliance. **Carmen is a Core Local Member of the Accenture Massachusetts Team.**

Denard Mondesir
Data Governance

Denard is an experienced strategist in IT and data transformation with a focus on public-sector modernization. He has delivered cloud adoption roadmaps, analytics programs, and change initiatives across large organizations, connecting business outcomes to technology choices. His background in program and financial management brings rigor to planning and execution. Denard is known for bringing structure and clarity to complex environments, helping teams adopt data and AI solutions with confidence and measurable impact. **Denard is a Core Local Member of the Accenture Massachusetts Team.**

Graciela Arana-Comas
Commonwealth AI

Graciela is an experienced strategist in AI value realization, specializing in Centers of Excellence and operating models for state government and higher education. She aligns CIO/CTO offices, business leaders, and delivery teams to prioritize use cases, govern responsibly, and move from pilot to repeatable execution. Her work spans health and human services as well as university settings. Graciela is known for practical playbooks that balance ambition with discipline—so programs scale without losing control. **Graciela is a Core Local Member of the Accenture Massachusetts Team.**

Emily Ferraro

Emily Ferraro is a Data Science Manager in Accenture's Data & AI practice with expertise in Generative AI strategy, intelligent workflows, and data-led transformation. She has guided clients across financial services, aerospace and defense, and health — designing AI Centers of Excellence, building multi-agent orchestration frameworks, and developing governance operating models. Emily is also a healthcare SME — having delivered precision health solutions, predictive churn models, and clinical dashboards that improve patient and member outcomes. Known for translating complex AI into scalable, high-impact solutions, she helps organizations

Agentic AI	realize measurable value through applied AI and advanced analytics. **Emily is a Core Local Member of the Accenture Massachusetts Team.**
 Dea Biancarelli **Training & Change**	Dea is a Management Consultant Manager within Accenture's Health & Public Service-Public practice with a passion for enabling communities and organizations to produce tangible impacts on today's pressing public health problems. Prior to Accenture, Dea worked in academic research with vulnerable populations and is highly skilled at translating and communicating complex research findings into persuasive real-world implications. Dea is most interested in utilizing her research and content expertise to effectively implement innovative strategies within government programs. **Dea is a Core Local Member of the Accenture Massachusetts Team.**
 Mimi Whitehouse **Responsible AI Governance**	Mimi is an experienced leader in data governance and Responsible AI with more than 15 years in health and public service. She has designed governance frameworks, model oversight, and workforce readiness programs aligned to policy and regulatory needs. Her work emphasizes practical controls, documentation, reviews, and monitoring that teams can run. Mimi is known for turning principles into everyday guardrails, enabling faster delivery without sacrificing trust or compliance. **Mimi is a Core Local Member of the Accenture Massachusetts Team.**
 Saugata Mandal **Platform**	Saugata is a seasoned technology leader in data and AI transformation with over two decades of enterprise experience. He has managed global delivery teams, led multi-country migrations, and modernized analytics on platforms such as Snowflake, Azure, and Salesforce. Saugata harmonizes complex legacy estates with modern cloud architecture, establishing reliable data pipelines and governance. He is recognized for achieving scale and measurable outcomes while keeping operations stable and cost transparent.
 Matt Johnson **Website UI/UX**	Matt is a seasoned service and interaction designer who turns complex systems into clear, accessible experiences. He combines research, co-design, and road-mapping—expertise he also teaches at the graduate level—to create interfaces and journeys that people understand the first time. Matt is recognized for applying design thinking in compliance-heavy settings, improving adoption, trust, and the everyday usability of tools and services.
 Cody Burns **Central Ledger**	Cody is a seasoned architect in distributed systems and digital assets, with experience across finance, health, and supply chain. He has designed award-recognized ledger and traceability solutions and contributed to industry standards. Cody is known for pragmatic architectures that balance security, privacy, and scalability—making multiparty systems dependable in production and translating novel technology into operational results.
	Logan is a Data & AI Security Lead specializing in cybersecurity strategy, AI governance, digital identity, and Secure AI. He has delivered large-scale information security transformations across both public and private sectors, helping organizations design and implement enterprise-wide governance frameworks. Logan combines cutting-edge AI innovation with practical implementation discipline, enabling clients to deploy ethical, resilient, and secure AI systems.

Key Project Personnel and Credentials

Beyond our core project team, you'll have access to Accenture's Luminary Network—renowned data, analytics, and AI experts who offer strategic advice and best practices, whenever you need them. Together, we're committed to building a solution that's secure, compliant, and designed for lasting impact. With the right people in place, we're confident we can help Massachusetts realize the full potential of the DCC.

Budget & Cost Proposal

At Accenture, we have deep experience delivering the technical program outlined in this proposal, and we have successfully delivered it on-time and on-budget. We also know how to deliver technology at scale in the Commonwealth and that these critical projects with limited state dollars must mitigate risk.

To provide MTC with a proposal that maximizes value and accountability, we are proposing a ***fixed-price approach*** to minimize risk of budget overruns and ensure timely high-quality delivery.

We have based our costs on our approach, assumptions and timeline outlined in the proposal and included in the attached **Budget Template (Attachment C)**

Key to our fixed price approach is a checkpoint at the end of month 1/discovery, where Accenture and MTC would mutually agree on an updated set of assumptions and technical and functional scope. Any implementation fee estimates will be refined as a result of the outcome of the Discovery Phase to reconcile changes and provide effort/cost for the full implementation. We view this as a key checkpoint to mitigate risk and increase accountability as part of the program.

As a part of our fixed-price approach, we have proposed an on-shore and off-shore staffing model that we have successfully executed for many clients, including the Commonwealth of Massachusetts. This approach centers on a Boston based team, led by our Executive Sponsor, Account lead, Program Manager and each Functional lead. They are supported by teams from Accenture Global Delivery Centers where they have the necessary expertise for each workstream. Data Security is key for this and many projects, and we ensure that we adhere to the data privacy and location principles laid out by the Commonwealth.

We would welcome the opportunity to discuss this approach for the foundation build that is included here as well as future phases and management.

Relevant Experience & References

Delivering Modern Data Management and Exchange for SINTEF (Independent research foundation, Norway)

Client Contact: Torbjørn Meistad, Enterprise Architect
Email: torbjorn.meistad@sintef.no| Phone: +4747335336
Project Scope: Implement a centralized, user-friendly data management and exchange platform for SINTEF's research institutes, supporting thousands of researchers and projects annually.
Project Description: SINTEF, one of Europe's largest independent research organizations, faced challenges with decentralized data management, duplication of effort, and lack of lifecycle enforcement. Accenture was selected through a competitive RFP to design, deliver, and operate a comprehensive solution that enables secure, accessible, and reusable data sharing across SINTEF's eight research institutes and 2,000 researchers, supporting over 35,000 projects per year.

Relevancy to MA: This project directly aligns with the Commonwealth of Massachusetts' priorities for people-centered digital services and cross-agency collaboration. The solution delivered for SINTEF demonstrates how a unified, accessible, and secure data management platform can break down silos, improve service delivery, and ensure compliance with data sharing standards—mirroring Massachusetts' vision for holistic, technology-enabled transformation in government-to-citizen digital services. The approach supports transparency, accessibility, and adaptability, which are core to the MA Digital Roadmap.

Description of Services:
To address the challenges of managing research data across multiple institutes, our team developed and launched a centralized platform that revolutionized how data was stored, accessed, and shared. This initiative focused on empowering researchers with secure, intuitive tools and ongoing support, while driving adoption through effective change management.

- Developed a centralized data management platform to unify research data across institutes.
- Enabled secure data sharing and enforced lifecycle policies to protect sensitive information.
- Designed user-friendly interfaces tailored for researchers, simplifying data upload, discovery, and sharing.
- Provided ongoing infrastructure and operational support to ensure platform reliability and scalability.
- Led change management and communications efforts to encourage adoption and facilitate a smooth transition.

Approach:
Our approach was rooted in collaboration and user-centric design, ensuring the solution met both institutional priorities and the practical needs of researchers.

- Collaborated with stakeholders to develop a robust business case aligned with institutional goals.
- Applied user-centered design principles, engaging researchers throughout the implementation process.
- Built the platform using an API-first strategy for flexible integration with existing and future systems.
- Incorporated policy-based controls and comprehensive audit features to ensure data security and compliance.
- Established a continuous improvement process, gathering user feedback and iteratively enhancing platform features.

Results:

- Centralized platform adopted across all institutes
- Improved data findability, accessibility, interoperability, and reusability
- Reduced duplication of effort
- Enhanced compliance with data sharing standards
- Ongoing feature enhancements

North Carolina Medicaid – AI Center of Excellence (CoE)

Client Contact: Sarah Gregosky, Chief Operating Officer
Email: Sarah.Gregosky@dhhs.nc.gov | Phone: 919-527-7027
Project Scope: North Carolina Medicaid, part of the NC Department of Health and Human Services, partnered with Accenture to launch an AI Center of Excellence (CoE) to accelerate responsible AI adoption across the agency. The initiative addressed the need for structured governance aligned with public sector values and regulatory standards.
Project Description: The CoE established a governance and risk management framework for AI and Generative AI to support compliance, transparency, and ethical use. Over three phases—research and design, validation and customization, and rollout—Accenture delivered foundational Responsible AI artifacts and operationalized governance processes.

Relevancy to MA: This engagement aligns with the DCC's emphasis on Responsible AI, governance, and risk management. The CoE demonstrates Accenture's ability to embed fairness, transparency, and compliance into AI initiatives, with the aim of engendering public trust and regulatory alignment.

Description of Services:

- Developed 15 foundational Responsible AI artifacts across governance, risk, and deployment readiness.
- Reviewed governance and compliance with regulatory requirements, leading to additional AI principles, AI policy, risk framework comprised of risk screening, risk tiering and risk assessment, and embedded procedures.
- Delivered training and change management for 18 cross-functional leaders.
- Established the AI CoE's operational oversight for AI initiatives across Medicaid programs.

Approach:

- Conducted stakeholder workshops to align governance with state priorities.
- Applied Accenture's Responsible AI framework for policy and risk management.
- Delivered iterative validation with legal, policy, compliance, and security teams.

Results:

- Established an AI governance model for the state's Medicaid agency.
- Enabled responsible design review of AI and Generative AI solutions.
- Ultimately, the AI CoE will support operational efficiency through structured oversight and risk mitigation.

California Department of Public Health (CDPH) – myCAvax Vaccine Management Platform

Client Contact: Maria Volk, Assistant Branch Chief, Immunization Branch at California Department of Public Health
Email: maria.volk3@cdph.ca.gov
Project Scope: In December 2020, immediately after the CDC granted emergency use authorization for COVID-19 vaccines, CDPH identified the need for a statewide vaccine management system. Accenture was selected as the prime contractor to design and implement myCAvax, leveraging Amazon Connect and Salesforce platforms to support vaccine distribution, provider enrollment, and dose accountability.
Project Description: Working alongside FEMA and Blue Shield of California, Accenture delivered a comprehensive vaccine management solution to enable rapid distribution and administration of vaccines across California. The platform supported provider onboarding, vaccine ordering, inventory management, and reporting, ensuring compliance with CDC requirements and minimizing waste.

Relevancy to MA: This project demonstrates Accenture's ability to rapidly design and deploy a secure, scalable, and user-friendly platform for critical public health needs—a capability directly relevant to the DCC's goals of interoperability, governance, and adoption. The solution's integration of data quality, analytics, and role-based access controls mirrors the Commonwealth's priorities for transparency, security, and equitable access.

Description of Services:

- Designed and launched myCAvax MVP in 3 weeks, using Agile delivery and cloud-native architecture.
- Enabled provider enrollment, vaccine ordering, and dose accountability workflows.
- Delivered Digital COVID-19 Vaccine Record (DCVR) and Digital Vaccine Record (DVR) for public access.
- Built vaccine analytics solutions for distribution, allocation, and forecasting across multiple programs.
- Implemented data quality solutions and real-time dashboards for equity and compliance reporting.
- Integrated HL7-based data exchange with immunization registries and enabled Master Person Index (MPI) functionality for accurate patient matching.
- Supported onboarding and training for over 26,000 users across providers, local health jurisdictions, and state agencies

Approach:

- Applied Agile SDLC with parallel development teams for rapid feature delivery.
- Built on Amazon Connect, Salesforce, MuleSoft ESB, and Azure/Snowflake/Databricks for data architecture.
- Established role-based access controls and HIPAA-compliant environments.
- Integrated advanced analytics and AI/ML forecasting for demand planning and equity metrics.
- Collaborated with CDPH and stakeholders through a robust governance structure to align on evolving public health priorities.

Results:

- MVP launched in 3 weeks; iterative enhancements delivered every 2 weeks.
- Scaled to support 26,000+ users including providers, local health jurisdictions, and state agencies.
- Delivered real-time dashboards for Governor's Office and public reporting within 60 days.
- Enabled seamless integration with California Immunization Registry (CAIR) and other public health systems.
- Achieved high system uptime and performance during peak vaccination periods.
- Enabled vaccination sites to deliver 8,000 doses/day per location.
- Supported 22 million vaccinations in 6 months.

Disclaimer

Legal Disclaimer

This document is proprietary and confidential to Accenture. It is provided in confidence and, except for evaluation purposes by the Commonwealth of Massachusetts and its designated agencies, may not be disclosed, duplicated, or used in whole or in part without Accenture's prior written consent.

The contents of this document are subject to formal contract negotiations. Nothing herein shall form the basis of any contract, and no part of this document will be binding on Accenture unless expressly agreed to under a formal contract. All representations and warranties, whether express or implied by statute, law, or otherwise, are hereby excluded.

For clarity, Accenture interprets the following terms as used in this solicitation and proposal: "maximize," "optimize," and "optimal" mean to improve to a commercially reasonable degree; "minimize" means to reduce to a commercially reasonable degree; "best" means leading or of a high standard; "partner" and "partnerships" refer to collaborative relationships, not legal partnerships; "right," where used as an adjective, means appropriate; and "ensure" and "enable" mean to use commercially reasonable efforts to implement.

This document was prepared based on instructions and information provided by the Commonwealth of Massachusetts and its agencies. Accordingly, Accenture accepts no responsibility for any inaccuracy, error, or action taken or not taken in reliance on this document.

These limitations are not intended to restrict ongoing business discussions between the Commonwealth and Accenture.

Assumptions

The following assumptions have been established to provide clarity and transparency regarding the scope, responsibilities, and dependencies associated with this proposal. These assumptions are foundational to our solution design, delivery approach, and effort estimates. They reflect our current understanding of the project requirements, client environment, and anticipated collaboration model. Any changes to these assumptions may impact on the proposed solution, timeline, or associated costs. We remain committed to working closely with the client to validate and refine these assumptions throughout the engagement, ensuring alignment and mutual agreement at each project phase.

Overall Solution Assumptions

- Infrastructure, Networking, Security
  - Accenture will be responsible for all AWS, Snowflake-related infrastructure configuration and sizing/ capacity planning.
  - Accenture will own all network and security changes required at the Infra level
  - MTC to ensure all the infrastructure – Environment, Servers are enabled by mid-November.
  - Assumed 3 environments dev, test, and production in scope for Data Engineering
- Location
  - Delivery Locations for this deal would be US, India, and Israel
- Source Data
  - It is assumed that public domain Source data would be available in S3 bucket. We will validate this during the discovery phase.
  - Data migration scope to be assessed as part of discovery phase.
  - MTC is responsible for provisioning test data and paying and contracting for data storage
- Ingestion
  - The source data would be ingested from S3 bucket to Bronze Zone of Snowflake Database using AWS Glue
  - The data will be Cleansed and Standardized while ingesting from Bronze Zone to Silver Zone using Snowsql
  - The data cleansing, Profiling and transformation would be done while moving data from Silver to Gold Layer using Snowsql and would be based on agreed upon data model.
  - ETL Pipelines
  - Same ETL pipelines would be used for History and Incremental Data load.
- Data Masking/Encryption
  - Data masking / Tokenization is out of scope. Default data at rest and transit encryption will be provided as part of AWS infrastructure.
- Stakeholders
  - Key stakeholders (decision makers, product owners, subject matter experts, etc.) from MTC business/IT units are available for workshops/clarifications/reviews as defined in the detailed project schedule.
  - MTC provides current state documentation as per requested timelines.
  - MTC's business users will be responsible for business validations. The client will provide required business inputs to support Accenture resources in performing the activities in scope. (i.e., client will assign PM, business analysts, SMAs, and architects as needed).
  - Workshops would be conducted to finalize the required KPIs for Gold Layer.
- Discovery / Planning Phase

- The Accenture team will be onboarded and granted the necessary platform access upon commencement of the engagement
    - MTC Source and Business SMEs should be available for workshops and discussion on scope, architecture and the implementation plan.
    - Product owners will provide business requirements and work jointly on data model design during discovery phase.
    - Semantic layer related requirements should be shared with the Data Modelers by mid-November, so that those can be used as the inputs for Data Modelling in Snowflake
- License
    - All licenses to be procured by MTC
- Stakeholder Management
    - Identification of key resources, SPOCs, BAs, for timely reviews, analysis deliverables/signoffs, source system owners & key decision makers
    - Identification of key parameters and data separation by release plan/phases

AWS Assumptions

- Cloud Foundation
    - The procurement / design and build of Direct Connect or other Network connectivity, AD, DNS from DCC to AWS would be established prior to the infrastructure being setup.
    - One primary region with Dev,QA,Prod environments and one disaster recovery region for Prod environment will be built.

Snowflake Assumptions

- Foundation Scope
    - Use selected priority Public data Sets
    - Leverage baseline Snowflake environment and feature set

Agentic AI Assumptions

- Building up to 8 Agents in AWS with Bedrock and leverage other AWS managed services as needed to support Ethical Upload use case.

Responsible AI Assumptions

- An initial set of metrics are agreed to and defined by the end of Month 2
- A quantitative definition of, or a dataset representing statistical fairness is available to perform bias review; otherwise, MTC will confirm quantitative measures of fairness
- Scope for fairness assessment is for a limited number of datasets to simplify feedback review
- Because custom models are assumed to be developed at a later stage, model assessment is out of scope for this phase
- MTC has identified individuals for the Fairness Oversight Board, and can designate individuals in the roles of fairness auditors, data stewards and reviewers. Formalization of the Fairness Oversight Board will take place in a future phase.

Credit Ledger Assumptions

- No fiat currency at launch; earn/spend only within defined credit economy

- Covers dataset version manifests, publish/share/access events, and verification flows only
- Only hashes and headers stored on chain; full payloads remain off-chain
- Initial implementation on chosen backend (QLDB/Fabric/anchoring) with pluggable design for future swap
- Emits version manifests (Merkle root) to ledger; no deep query rewrite in Foundation Release
- Capture Zero Copy/Reader grants and access approvals: no advanced marketplace monetization
- KMS-backed key lifecycle; no external identity federation in Foundation Release
- Basic earn/spend rules, issuance caps, throttles, expirations; sustainability incentives included but no dynamic pricing
- One-click "verify dataset version/access" with proof trail; Admin UX limited to credit issuance and abuse review
- Core metrics (cost per 1k events, issuance vs burn, sustainability) and dashboards; no advanced anomaly detection
- STRIDE/LINDDUN threats addressed; RTBF via pointer rotation; RAI review for agent interactions
- Scale target defined for Foundation Release (events/month) with confirmation latency and availability SLOs; not full production scale

Security Assumptions

- Accenture will utilize Accenture Third Party Software licenses for the mySecurity Diagnostic service from Securiti.ai (SaaS product).
- Securiti.ai will retain ownership of its Intellectual Property Rights and Materials, which will be considered as Accenture Third Party Materials.
- Accenture will utilize Securiti.ai for asset cataloging and data discovery scanning of one data store. The solution will be SaaS based and only one scanner (pod) will be deployed within MTC production environment.
- Accenture is not licensed or certified in any country, state, or province as a public accountant, auditor or legal advisor, or private investigator and is not being retained to provide accounting services, accounting guidance, audit or internal control advisory services, tax or legal advice or investigatory services that would require a license.
- Accenture and Accenture business partners (Security.ai) may review and analyze MTC-specific metadata gathered through the appraisal process for the purpose of preparing the diagnostic reports.
- The infrastructure provided by MTC will meet Securiti.ai technical requirements such as ability to create service accounts, apply necessary roles to service accounts, and provisioning of service accounts to in-scope cloud environments.
- MTC agrees to the use and scanning of a data store in their environment in read-only mode, and ingestion and temporary storage of cloud resource metadata for analysis.
- MTC provides proper resources to install a Securiti.ai scan pod in MTC environment on a VM per vendor requirements and connect to the selected data source.
- MTC understands and acknowledges that Accenture will use third party technology (Securiti.ai) to support Accenture's performance of the Services.
- MTC will provide resources to support the onboarding (which consists of granting access to Securiti.ai application) of in-scope data store within project timelines.

Website and Admin UI Assumptions

- Staffing is based on the assumptions made on the number of screens and their complexities. Discovery phase will be used to validate these assumptions, and the staffing plan will be trued up based on discovery outcomes.
- All UI/UX designs need to be approved by MTC before the frontend team can start the implementation. Any delay in reviews will affect timeline and cost.
- For Foundation Release, only English language is supported though the app will be architected to support multiple languages.

General Assumptions

- MTC will identify and commit the necessary personnel, management involvement, access, and quantitative information to support the Project, including the participation of individuals in working sessions and meetings.
- MTC will review and accept deliverables and make project-related decisions in a mutually agreed timeframe consistent with the project plan outlined in this document.
- MTC will provide access to all reasonably required information, data, documentation, and other items needed to complete the project accurately within the agreed timeframe.
- MTC shall be responsible for obtaining, at no cost to Accenture, consents for Accenture's use of any third-party products, including, but not limited to software (including purchase of any licenses), necessary for Accenture to perform its obligations under this SOW.
- MTC shall assist with and facilitate coordinating and scheduling meetings, including setting up of web-conferences and reserving conference rooms or meeting space required where Accenture personnel do not have direct access to the resources themselves.
- MTC shall be responsible for its operation and use of the deliverables and for determining whether to use or refrain from using any recommendation that may be made by Accenture. MTC will be solely responsible for determining whether any services provided by Accenture (i) meet the MTC's requirements; (ii) comply with all laws and regulations applicable to MTC; and (iii) comply with MTC's applicable internal guidelines and any other agreements it has with third parties.
- Each Party (or its licensors as applicable) shall retain ownership of its intellectual property rights, including without limitation patents, copyright, know-how, trade secrets and other proprietary rights ("IP") which were existing prior to each respective Statement of Work ("SOW"), or IP developed, licensed or acquired by or on behalf of a Party or its licensors independently from the Services or the Deliverables, in each case including any modifications or derivatives which may be created as part of the Services (collectively "Pre-Existing IP").
- MTC shall be responsible for the performance of other contractors or vendors engaged by MTC in connection with the Project and ensuring that they cooperate with Accenture
- MTC directed changes to project scope, approach, timeline, and deliverables may impact the schedule, fees, costs or level of effort. Other than an amendment to the contract, Accenture will provide written notification to the state of any such impact within three business days of a change affecting the scope, approach, timeline, deliverables and/or cost.
- Accenture in providing services is not licensed to provide and will not provide any legal advice of any kind, whether governed by law or custom, in connection with or as part of the services.

Deliverables

Accenture will deliver the technical foundation the AI Data Commons needs to fulfill its promise. Our integrated platform architecture, agentic AI framework, and secure data foundation ensure

the system performs reliably, scales responsibly, and earns trust from day one of the Foundation Release.

Every deliverable, from design documentation and implementation scripts to fairness reporting, security plans, and codebases reflects our commitment to building systems that work for people, not just process requirements. We combine proven methodologies with deep public-sector experience to ensure every solution domain is addressed with purpose and precision.

With Accenture's global capabilities and Massachusetts-rooted expertise, the Commonwealth can expect deliverables that don't just meet standards, they establish new ones for responsible AI, inclusive innovation, and public value.

SR #	Solution Domains	Key Solution Components	Deliverables
1	Overall Solution	Integrated Platform Architecture	Design Document
2	Agentic AI	Agentic Framework	Agents' Design Documentation
3	Agentic AI	Agentic Framework	Agents' Implementation Scripts
4	Data Foundation	Data pipeline, Catalog& Metadata Interface	Design Document
5	Data Governance	Ethical Access, Fairness and bias, Human in the Loop	Fairness and bias tools and reporting
6	User Interface	UX Design, Credit Ledger Design	Design Document
7	User Interface	Portal	Codebase
8	User Interface	Credit Ledger	Codebase
9	Credit Ledger	Credits Economics, Observability	Design Document + Configuration
10	Security	Data Security, Compliance	System Security Plan (SSP) with the documented controls
11	Security	Application Security	Application Security Documentation - test results, issue and remediation log
12	Security	Infrastructure, Platform Security	AWS Security (Native Services) design document Integrations documentation (SIEM & CSPM)
13	Training	Training Curriculum	Persona-based curriculum, quick-start guides, micro-lessons, hands-on labs, help center content
14	Operational Readiness	Manuals/Job Aids	Internal runbooks (AWS, Snowflake, Credits ledger), governance templates, security patterns
14	Engagement	Stakeholders Reports and Strategy	Stakeholder discovery report, Change Champion Network charter, co-development proposals
15	Communications	Materials	Outreach campaign assets, media content, event toolkits